Embracing love and friendship with robots

An anthropological and philosophical analysis of intimate human-robot relationships

Alice Yunyun Wang

Embracing love and friendship with robots

An anthropological and philosophical analysis of intimate human-robot relationships

Alice Yunyun Wang

Master Thesis

Supervisor: Dr. Julia S. Hermann
Second reader: Dr. Naomi Jacobs

MSc Philosophy of Science, Technology, and Society
Faculty of Behavioral, Management, Social Sciences
University of Twente, Enschede, the Netherlands

November 2022

Table of contents

“Love isn't something natural. Rather it requires discipline, concentration, patience, faith, and the overcoming of narcissism. It isn't a feeling, it is a practice.”

Erich Fromm, *The Art of Loving*

Abstract

This thesis presents a philosophical-anthropological study of intimate human-robot relationships. The thesis aims to better understand love and friendship between robots by putting different perspectives together and examining where they differ and what they have in common. Conceptual analysis is chosen as the method to investigate the research question: how can we make sense of the phenomenon of romantic love and friendship in human-robot relationships and is it something we should strive for? I explore the notion of love in human-human relationships and the anthropological-psychological explanations for human-robot relationships, analysing love's epistemological and moral values, and reasons why humans may develop affective feelings toward robots. I outline the necessary conditions for mutuality to be achieved in intimate human-robot relationships and examine the possibilities of artificial consciousness for robots to achieve mutuality. Anthropomorphism demonstrates humans' longing for social connection and self-recognition, while the Uncanny Valley and Dehumanisation hypotheses reveal the desire for grasping the notion of the human self. Given love's uniqueness and its moral and epistemic value, I argue that intimate human-robot relationships authentically exist and should be advocated and encouraged as a tool to cultivate compassion, self-transcendence, and social progress. The love that may occur between humans and robots is non-individualistic and should not be judged but respected. Engaging in loving relationships with robots can help to expand human's ability to socialise, cultivating compassion and love in more meaningful and ethical ways.

Introduction

Nowadays, humanoid robots are widely used in society. Humanoid robots operate in the fields of personal assistance, caregiving, education and entertainment sections, and they take on roles to achieve whatever they are designed to perform. Recent technological advancements allow humanoid robots to be produced to highly authentic levels (i.e., humanlike), where social robots are able to display emotions and anthropomorphise human behaviours in order to realistically mimic human interactions. As a consequence, some users have developed an emotional affinity towards individual robots. The users may express affective feelings and some even claim to be in love with their humanoid companions. More surprisingly, users may not only become involved with robots in order to fulfill their physical desires, but emotional bonds appear to be existing as well. Customers of sexbots hope for manufacturers to produce robot partners with the ability to think. Some users have difficulty maintaining interpersonal relations and employ sexbots as an effective tool to practice or engage in intimacy (Castello, 2018). However, while robot emotion expressions are deemed commercially valuable and philosophically interesting, the understanding of emotions in human-robot interaction remains largely underexplored (Chuan & Yu, 2021).

Humans tend to express empathy towards objects. In psychology, anthropomorphism is the innate human tendency to attribute human characteristics to inanimate objects. For example, despite looking nothing like humans, battlefield machines can trigger emotional responses (or even attachments) to humans working closely with them. Soldiers formed bonds with war robots, as they experience anger and loss when fighter robots are destroyed in combat (Neal, 2013). More intimate bonds seem to form more easily when robots resemble more human appearances. One more well-known case is the Chinese AI engineer Zhang Jia Jia who built his wife in his apartment. Zhang plans to teach his wife to walk and do chores around the house (Huang, 2017). Interestingly, the uncanny valley theory has demonstrated that humans positively experience robots that resemble

themselves, but revulsion and eerie feelings arise when robots become too humanlike (Mori, 2017).

The above cases of human-robot connections raise exciting philosophical and anthropological questions, such as 1) What is the nature of love? An investigation into the notion of love provides a better understanding of the phenomenon of love and friendship between robots and humans because it opens up the 'black box' of why humans love, subsequently shedding light on the possible reasons why humans love humanoid robots. 2) What do our ambivalent emotions towards realistic humanlike robots disclose about ourselves as humans? Answering this question yields knowledge on the formation of the contemporary human selves situated in a world in which technologies intrude on the very fabric of human selfhood. 3) What does the willingness and acceptance to bond with robots demonstrate in our current societies? Such a question explores the ways one understands and reflects robot developments through the lenses of subject-object dualism. And more fundamentally, 4) from an anthropological perspective, what (if any) demarcates humans from highly humanlike robots? This question is fundamental to the understanding of selfhood and self-formation, whereby yielding what distinguishes a robot and a human adds insights to the conceptualisation of human beings. Together, these questions contribute to the main research question of this thesis: **How can we make sense of the phenomenon of romantic love and friendship in human-robot relationships and is it something we should strive for?** The motivation to answer this question is to obtain a better understanding of intimate human-robot relationships by collectively putting different perspectives together, and looking into what these perspectives have in common and how they differ.

The motivation for this thesis topic is twofold. Firstly, the notion of love is considerably one of the most complex and elusive concepts to philosophically define and analyse and this contributes significantly to how love is expanded to technological artefacts through human-robot interactions. Secondly, I find the implications of designing robotics in

anthropomorphic ways intellectually fascinating, as it invites me to think about how designing artificial emotions allow users to develop affective feelings about technologies that merely resemble human appearances. I choose to write this thesis in a predominantly descriptive manner, instead of a normative one. This is because most existing philosophical literature on the love and sex between humans and robots ground themselves in different normative stances and advocate certain directions of interacting with robots. I find writing a descriptive essay such as this one unique and conceptually valuable in this case. Having said that, this present thesis also contains a minor normative stance, which is that intimate human-robot relationships facilitate positive outcomes in love and friendship and therefore is morally desirable and should be encouraged.

In order to answer the main research question about the phenomenon of love and friendship in human-robot relationships, this thesis is structured into three chapters. The first chapter explores the philosophy of love. In particular, the nature of love, the moral and epistemological value that love brings about, love's dual nature, love as a psychosocial condition, and the biological dimension of love. The first chapter lays a foundation for the next chapters. By first obtaining a certain comprehension of human love, the second chapter adds robots into the equation. The second chapter discusses possible determinants for humans to seek robot companions from psychological and anthropological perspectives. Following that, a brief history of social interaction between humans and computers is presented in order to illustrate how such social interactions came about through historical circumstances. The meaning of gender in robots is discussed afterwards, questioning the roles of sex and gender that arguably are both applicable to humans and robots. The following sub-questions will be answered throughout the chapters:

1.
    a. What is the nature of love?[1]

[1] Note: Although friendship is not explicitly stated here, love between friends are no less of love than romantic love, to my understanding.

b. How does love relate to morality and epistemology?

2.

a. Why do some humans desire intimacy with robots rather than other humans?
b. What does the obsession with humanoid robots disclose about ourselves?

3.

a. How can a robot love a human mutually and to an Aristotelian ideal?
b. What social values do genuine human-robot relationships bring about?

Objective

This thesis aims to provide an in-depth understanding of the decreasingly absurd 'interpersonal', social dynamics that are increasingly prevailing between humans and robots. More specifically, this thesis is concerned with the themes of romantic love and friendship — which are typically categorised as human relationships. To my knowledge, there is a lack of aggregated philosophical papers that address the underlying processes, degree of feasibility and societal implications of the emergence of intimate human-robot relationships. These issues should be more cohesively explored and explained in philosophical manners. The central objective of this thesis is to understand the emerging phenomenon of human-robot intimacy in a digital age from anthropological perspectives. As a secondary objective, the thesis examines the current potentials and possibilities to achieve the notion of mutuality in human-robot relationships that are directed toward romantic love and friendship. Philosophical-anthropological arguments are primarily adopted by this thesis because it strives to understand human cultures and societies and the development thereof. Therefore, I find it most appropriate for anthropological approaches to fit the theme of this thesis, as it analyses how humans are evolving in the technological age.

Methodology

This thesis analyses users' emotional bonds and attractions to their robot companions. I plan to do this by first researching extensively into the (human) concept of reciprocal love, and afterwards applying and relating love to human-robot interactions. I will be attempting to explain how and why love can be an emotion to occur in human-robot relations. The goal of this present thesis is to get a better understanding of the phenomenon of love and friendship between humans and robots by collectively putting different philosophical perspectives together. My underlying assumption is that the phenomenon in focus here is a phenomenon that is facilitated by the current technological environment, therefore, specific theories and ideas belonging to the philosophy of technology and philosophical anthropology of technology are selected to put into this thesis. This thesis builds its arguments by means of a literature review and conceptual analysis. A literature review provides summaries of relevant research that took place prior to this present thesis. Conceptual analysis, on the other hand, seeks to understand the necessary and sufficient conditions of the being of the topic or subject in focus, (i.e., what are the necessary and sufficient conditions that make up an intimate human-robot relationship). Conceptual analysis is understood as an analysis of certain concepts, assertions, hypotheses, and theories (Petocz, 2010). Conceptual analysis looks for revelatory definitions (in contrast to analytical and stipulative definitions), whereby revelatory definitions use "perfect tendencies of thought implicit in old usages, offering more insight into the subject matter being treated" (Sasser, 2019, Timestamp 2:19). For this reason, this thesis selects relevant classical theories and ideas on love and friendship, such as Aristotle's virtue friendship, as well as relevant philosophical ideas of humanly love in literature, such as Jollimore's vision view. Subsequently, I apply these theories to the phenomenon of intimate human-robot relationships in order to conceptually analyse the essence of it, and find out the necessary and sufficient conditions for such relationships in a technological world.

By selectively reviewing and critically reflecting on the philosophical ideas pertinent to the topic of intimate human-robot relationships[2], I arrive at my main argument and the answer to my research question. Here, perhaps it is helpful to clarify why I choose to exclude the focus of a sexual component in this thesis. The reason is that there already exists extensive philosophical literature on sex robots and on humans engaging in sexual acts with robots, both in descriptive and normative fashion (Benlo, 2016; González-González et al., 2020; Szczuka & Krämer, 2016). In my view, adding to this existing literature is less important and interesting than the exploration of the love realm between humans and robots. Moreover, I personally perceive sexbots as not much different than an elaborated, more sophisticated version of regular sextoys, which forms another reason for leaving out the sexual aspect in this thesis and focusing on the themes of love and friendship instead. That being said, sex is also an essential aspect of romantic love, therefore, without a scrutinisation of it in this thesis, the sexual component is still mentioned, but merely as a stepping stone to come to a better understanding of love and friendship between humans and robots. Moreover, I am aware that some readers may be skeptical about the scope of this thesis. As it is likely for one to view romance and friendship as distinct topics to be focused on in one thesis, I would like to defend the scope of this thesis. The theme of this thesis encompasses the subject of love, which is most prevalent in romantic love and friendship. The thesis explores the abstract notion of love which consists of both romantic love and friendship in relation to robots and humans, instead of focusing on one particular type of love to analyse the phenomenon.

Further regarding the choices of methodology, I adopt theories and hypotheses from various philosophical disciplines with the goal to construct a more holistic picture put together from different philosophical perspectives. As such, this thesis arrives at its conclusion through the enlightenment of philosophies on love, humans and robots grounded in different branches. To illustrate, I adopt the ideas of Troy Jollimore who adds

[2] Whether the literature reviewed in this thesis is pertinent or not, it is of course according to my own judgement.

ethical and epistemological insights to love, this is to explore what values are embedded in the act of love, thereby revealing potential positive things that there could be when humans engage in intimate, genuine relationships with robots. I also adopt ideas found in the book Exitmate technology written by Ciano Aydin, who gathers relevant philosophical-anthropological knowledge and presents some explanations for the Uncanny Valley phenomenon. I found this to be coherent to be included in this thesis as his arguments to the Uncanny valley theory prescribe a logical and sophisticated answer to a timely mystery of the dynamics between many robots and humans (see section two). Thus, a better understanding of why highly realistic humanoid robots induce feelings of uncanniness and fear helps this thesis to conceptually analyse other human feelings towards robots apart from the positive ones. I believe that this adds depth to the research of intimate human-robot relationships, since the negative feelings towards robots may reflect and reveal more explanations of the positive feelings towards robots that are central to this thesis.

Moreover, I include the ideas of artificial consciousness from philosophers of mind such as John Searle, as artificial consciousness helps better understand the question: how to philosophically define a human being, considering that one day we may duplicate human consciousness? In addition, the thesis also includes knowledge from the field of psychology of biology to inquire about the essence of love and how it relates to intimate human-robot dynamics. I choose to do so because first of all, biology (which to a large extent includes psychology) is the most obvious difference between a human and a robot, thus understanding the human biology of why we love is foundational to move on to understand why a human and a robot may fall in love or become a genuine friend. Second, love is either exclusively biological or social, it is simultaneously both at the same time, and this will be discussed in detail in the first chapter.

Chapter one

Love and friendship in human-human relationships

This chapter discusses the inexhaustible topic of love within human relationships with a particular focus on friendship and romantic partnership. The objective is to provide a foundation for investigating the similarities and differences of intimacy and love within human-human relationships and human-robot/technology relationships. I review and discuss some ideas of the philosophy of love, with specific themes on epistemology, morality, (anti) rationalism, and normativism. The following questions are addressed in this chapter: 1) How does love relate to morality and epistemology? 2) What is the dual nature of love? 3) What is the biology of love and to what extent does it define love? And 4) Why should we view love in the light of non-individualism? Afterwards, this chapter presents the biochemical dimensions of the notion of love in order to understand love from a scientific perspective, in addition to social, psychological, and philosophical ones. Accordingly, research into the nature of love in human relationships also builds up knowledge for exploring the sub-question of whether mutual love is in fact possible between humans and machines.

Love is essential for a good life. Having positive interpersonal relationships produces many benefits, including pleasure, joy, and improved overall well-being (Gheaus, 2018). In the absence of loving relationships, humans tend to feel empty and depressed, even if all their basic needs are met and as such, loving relationships are vital for human flourishing (Carters & Porges, 2013). In terms of the depths of human relationships, friendship and romance are considered intimate, and one key element in loving relationships is reciprocity. It is not controversial to claim that mutual love is inevitably present in all genuine interpersonal relationships, but the notion of love and its nature is still elusive and interpreted in many ways in both academia and popular cultures.

As it has already been said in the introduction, despite sexual attraction being a necessity in romantic love, this thesis deliberately chooses not to focus on it, as there is an abundance of literature focused on sex in romantic relationships. That being said, if the weight of sex is minimised in focus, to my understanding, genuine, romantic partnerships and friendships are not so different because wholehearted love exists in both. Perhaps, the most notable difference is that, conventionally, in romantic relationships, love is exclusively shared between two people, whereas friendship is more acceptable by society to be shared by two or more people. However, free love and multiple partners have increasingly popularised when it comes to romance (Douthat, 2015; Klein, 2021; Vogels & Anderson, 2020), it is no longer deemed unethical or adulterous, or at least it is no longer a social taboo (Iturriaga & Saguy, 2017; Newport, 2020;).[3] Just like same-sex marriage, multi-partner relationships are increasingly socially, morally and legally accepted (Newport, 2020; Scaringi Law, 2022).[4] Moreover, in many of popular beliefs, the ideal romantic partner is also the best friend of the other, and many romantic partnerships initiated from friendships (Stinson, Cameron, & Hoplock, 2021). In addition, in romantic relationships, a sexual component exists inevitably. Nonetheless, this chapter includes some explanations of the reason for sexual attraction but it is beyond its scope to scrutinise the element of sex in love. This chapter focuses on an ideal of love, the type of love that is selfless or not egocentric in the least sense. Naturally, some elements may still differ (e.g., monogamous romance), but I think they are irrelevant to the research question this thesis attempts to answer. The rest of the chapter refers to love and friendship in

[3] This is also a dominant position commonly defended by feminist philosophers and sociologists (Iturriaga & Saguy, 2017).

[4] Newport (2020) provides research on the growing views on the acceptance of Polygamy in the United States. In 2004, around 7% of Americans found polygamy morally acceptable, whereas by 2020 the view that polygamy is morally acceptable has increased to 20% (Newport, 2020). In the Netherlands, although marrying more than one individual is legally prohibited, a samenlevingscontract is recognised as legal between more than two partners (Geysegom, 1997). A samenlevingscontract is "a notarial deed in which two (or more) persons that live together agree to take care of each other financially", it is translated to a cohabitation contract in English (Infotaris, 2019, p. 1). It is considerable to note that my point of this claim is that though many countries still legally prohibit polygamy, it is increasingly morally acceptable in the world (see appendix A). Moreover, in my view, love should not be exclusively legally defined by the act of marriage.

intimate human-to-human relationships as simply 'love', since conventional understandings of love are exclusive to humans.

Love as something in-between

Within the field of Philosophy of Love, there has been a revival of interest in the nature of love in recent decades (Jollimore, 2011; Mckeever, 2012). In particular, many philosophers of love sought to explain and rationalise the nature of love and attempted to explain the selectivity of love, i.e. why we love particular people. According to McKeever (2012), there have been two opposing schools of thought accounting for this topic, namely rationalism and anti-rationalism. Using one's normative intuitions, the former contends that love belongs to the category of emotions, and like other kinds of emotions, love occurs with reasons grounded in it and therefore can be justified and explained. To exemplify, oftentimes people appeal to the attractive or admirable qualities of the persons they love, to which they justify their love by citing different reasons. Similarly, the fittingness of instances of love is too commonly understood through judgement-making in attempts to construct some sort of logical explanation. These explanations in turn position the adequacy of such fittingness and assume the need to be backed up with reasons (Jollimore, 2011). For this reason, rationalism holds that, just like beliefs and emotions, love is an attitude which can be contrived fitting by reasons (Han, 2021). On the other hand, anti-rationalism argues that there are no justificatory reasons for loving particular people. Indeed, love sometimes appears to be oblivious to reasons. For example, it would be absurd and nonsensical to be unjustified in continuing to love the ones we love, even if they lose the qualities to which we appealed in the first place (Frankfurt, 2009; Han, 2021). The antirationalism of love is commonly represented by the phrase 'love is blind'.

Jollimore (2011) accommodates the rationalist idea that love is guided by reasons, as oftentimes the lovers can confidently provide reasons for loving the people that they do, yet Jollimore also points out that love is, at the same time, frequently associated with irrational and uncontrollable characters. Nonetheless, Jollimore (2011) points out flaws in

both rationalism and anti-rationalism to put forth a third account of love, he argues that love in fact is both guided by reason and blind. As Jollimore (2011) explains:

> *Part of the key to understanding the rationality of love is to avoid misunderstanding the ways reasons work. In particular, it is necessary to understand the way in which values generate reasons that render certain objects or opinions eligible for choice without making it mandatory that agents select those particular opinions. Such a conception of reasons leaves an agent free to recognise the existence of certain values and value bearers without being rationally required or compelled always to respond to them in the fullest sense. She may, that is, judge something to be valuable without valuing it herself. This is what makes it possible for a person to love without contravening the requirements of rationality.* (p. 93-94)

In the quote above, Jollimore argues that reasons do not fundamentally clash with irrationality. He emphasises the perspectives of the lover, where one can present reasons why they love others, regardless of whether the reasons are irrationally justified or otherwise, they are still the reasons that drive the lovers to love the ones they do. In other words, blindness and rationality co-exist in love. A similar idea of love can be found in the Symposium when Socrates discusses the nature of Love with Diotima (Plato, trans, 1989). Diotima reveals that Love is something in between (Plato, trans, 1989, 203a). It is neither good nor beautiful because Love needs things that are good and beautiful, yet this does not mean that Love is bad and ugly. And that Love is neither mortal nor immortal rather, it is a spirit that falls in between man and God. In the end, the dialogue leaves its listeners with the message that the typically grandiose, perfect picture of love is actually directed at the object of love and not the ones who love (202d). All the great qualities of love are what the lover seeks and thus lacks (Plato, trans, 1989, 202d).

Love's epistemic and moral values

On a similar note, for Jollimore (2011), love is also not absolute, for it contains both epistemic and moral quandaries, which he shows through his observation in the shoes of a typical lover who sees her beloved in the best possible light:

> *She's kept her love for him as alive as the summer they first met. In order to do this, she's turned life away… Once Uncle Julian told me how the sculptor and painter Alberto Giacometti said that sometimes just to paint a head you have to give up the whole figure. To paint a leaf, you have to sacrifice the whole landscape. It might seem like you're limiting yourself at first, but after a while you realise that having a quarter-of-an-inch of something you have a better chance of holding on to a certain feeling of the universe than if you pretended to be doing the whole sky.* (Krauss, 2005, as cited in Jollimore, 2011, p. 28)

Accordingly, Jollimore speculates that love is said to be "blind" due to several reasons, in addition to his judgement that love is a sort of vision belonging to the ones who love. Jollimore backs up this idea: "Among human beings, only the existence of those we love is fully recognised" (Weil, 1952, as cited in Jollimore, 2011, p. 88). First, it is epistemically problematic that love demands the lover to embrace, or at least to accept, a disproportionately, perhaps unjustly favourable view of the beloved. As such, the lover tends to idealise the beloved and subsequently, the lover is blinded to the negative qualities of her beloved. Jollimore speaks of love in terms of vision, as he refers to this account of love as "the vision view" (2011, p. 28). The vision view holds that love is fundamentally a method of perceiving (Jollimore, 2011, p. 88). The vision view acknowledges the truth in the saying 'love is blind', yet argues that love being guided by reason is not conflicted with or undermined by the blindness of lovers: "Reasons play an important role in love even if they rest on facts that are not themselves rationally justifiable" (Jollimore, 2011, p. 34). To define love as a way of seeing the world reveals salient implications for the nature of love. In no specific order of importance, love then is strongly associated with the following terms: limits, particularities, and choices/selectivity. These keywords characterise love's dual (anti)rational characters. Moreover, love's blinding property also illustrates that love does not solely belong to the rationalist school of thought.

Importantly, Jollimore identifies the difference between love and infatuation and stresses the danger of mixing them, since when one is infatuated with another, he is way more prone to idealise the other, producing a false image of the "beloved" in which he puts

himself, and results in a perfect bubble for the lover that will inevitably break someday only to leave him to a confusing and possibly bitter reality (p. 123). Nonetheless, it is arguably healthy to admit to ourselves that when we love, some positive biases are present, for they are natural effects of the act of loving. However, in order to see the beloved in more accurate ways than we would if we were not to love them, Jollimore (2011) indicates that the lover is required to obtain in-depth knowledge of her beloved in addition to understanding the beloved's reasons behind her actions (p. 128). Jollimore writes: "To love a person is to treat him as an end in himself and to fully recognise his existence as an individual" (2011, p. 167). Obtaining knowledge is desirable because the lover will not only love the beloved's body, but also her character traits, values, principles, and the overarching personality that makes the beloved the lover's beloved.

Secondly, morally speaking, a problem lies in the fact that love has the tendency to nudge the lover to avoid acting and reacting to the needs and wants of others in favour of her beloved, regardless of when the needs are greater in the others who are not the beloved of the lover (Jollimore, p. 29). This way, a tunnel vision overcomes the lover's rationality and objectivity that exist also outside of the domain of love, closing off herself from those who desire and need apart from her beloved, and to a certain extent ending up neglecting the morality of love. The morality of love refers to love's moral status. Jollimore (2011) argues that love, at its basic, is a moral phenomenon. For Jollimore, love is a deeply moral emotion that demands the lover's empathy, to take the beloved's concerns as one's own, which prevents excessive self-concern. At the same time, the passion that is inspired by love can cloud one's judgement and occasion in greatly immoral actions, it can blind one to the needs of other people than the ones he loves (Jollimore, 2011). As discussed in the beginning of this section, love is something in between. Hence, Jollimore reveals that love is neither absolutely moral or greatly immoral, neither wholly rational nor deeply irrational. Interestingly and surprisingly, even though the tunnel vision of love may lead us to look the other way when others' needs are calling for us, it creates an authentic and novel way to understand love's morality. The preconditions to understanding the morality of love lie in

the effort to achieve a full appreciation of someone else in their uniqueness and individuality as well as to comprehend her feelings, thoughts and actions in the most charitable light (p. 47). When these preconditions are met, love resembles an ideal moral relationship (Jollimore, 2011). Relating to this, Jollimore (2011) further points out that love provides lovers and those who once loved with insight into the value of other people, and is thus capable of teaching them to be moral (p. 146). As such, we may ask ourselves: what is the relationship between love and morality? Jollimore (2011) claims that morality is only made possible through the existence of love, in other words, the two are intricately related. He suggests that one could be moral even without love when the impersonal moral rules are followed, however, the said impersonal moral rules are only able to be developed in the first place via identifying with others and fully appreciating their intrinsic value: "the impersonal saint who lives her life entirely according to the impersonal attitude could only know that others are valuable through the experiences of lovers" (Jollimore, 2011, p. 168).

> *The impersonal saint knows that human beings are valuable, but how does she know that? If her relations to others are genuinely and pervasively impersonal, she cannot know it through direct experience, for direct experience of that value, or of any value, is always particular. She would have to know it, then, through the experiences of others: the experiences, that is, of lovers.* (Jollimore, 2011, p. 168)

The biological dimensions of love

Love is not only a feeling. It is also something deeply biochemical. Love is a dynamic and bidirectional biological process that pervades every aspect of a human's life (Cater & Porges, 2013). Within loving relationships, the maintenance of positive, rich interpersonal dynamics requires continuous feedback via the sensory and cognitive systems; the body constantly seeks love and the body correspondingly responds continuously to interactions with the ones it feels affection towards, or to the lack of such social behaviours and interpersonal engagements (Carter & Porges, 2013). Take social interactions as an example, meeting and talking to others in a social situation stimulates the cognitive and physiological processes and in turn, the emotional and physical states are influenced by it.

Subsequently, these triggered changes in states and influences will affect future social interactions. As such, one's physical and mental states are profoundly affected by love, whereby a "broken heart" can be of calamitous psychological consequences and negative emotions such as bereavement sabotage the human physiology by accumulating stress and depressive thoughts, which in some cases can even be fatal (Carter & Porges, 2013).

So what is the biology of love? The answer to this question to a significant extent, if not completely, concerns evolution. In fact, as the renowned biologist Theodosius Dobzhansky once famously wrote: "nothing makes sense in biology except in the light of evolution" (Dobzhansky, 1973). Such a rather bold statement accounts for the biology of love too. Humans, like all sentient organisms, live lives that are essentially social. However counterintuitive this statement may be for attempting to explain the biology of love, growth, reproduction and mutual homeostasis are fundamentally supported through the organism's ability to interact and communicate with others dynamically. Even long before the existence of humans, social engagement and interactions were already discovered to be among primitive invertebrates (Carter & Porges, 2013). Single-cell organisms, bacteria are able to identify and approach the fellow members of each of their own kinds. Moreover, the reproduction of bacteria has also a higher success rate when they are in the company of their own species and when they are able to form communities with chemical and physical particularities that go far beyond what an individual cell is capable of (Cater & Porges, 2013). Thus far, one may conclude that at its essence, love from the biological perspective entails a collectivist tone, where life thrives through love when companionship is present, and life is more prone to be defeated or fail when one is alone whereby love can be understood as absent.

Despite there being diverse theories on the topic of love in contemporary neurobiological and social sciences, there is a prominent account that attempts to capture the essence of love in organic beings (i.e., humans and other mammals) by splitting it into three primary emotion systems (Wu, 2017). These are the lust, attraction and attachment systems, and

they can all be explained by fluctuations in different hormone levels within the human brain. Since the primary focus of this thesis is love in relationships, the latter two emotional systems would be more in-depth discussion, as the lust system is more related to sexual intents and less concerned with love stripped down to its bare bone. However, the lust system is not irrelevant since, within human-robot intimate relationships, half of the party still would be driven by their libido to pursue love within such relationships. These three brain systems have evolved to suit the social and reproductive needs of modern humans, they function distinctively but overlap when it comes to being in romantic love and intimate platonic love (Wu, 2017). The former two systems are almost solely exclusive to romantic relationships for their mating, reproductive and parenting purposes. In contrast, the latter also holds true of its existence in meaningful, well-developed and maintained friendships. Assigned by natural selection, the lust system resembles one's libido, and its role is to inspire and facilitate desires in a range of potential mating partners (Wu, 2017).

I choose to include the biological explanations for love as seen above because most notably, biology demarcates something that is alive from non-living things, which also seems to be the biggest difference between a human and a robot. However, as technology advances through time, such a demarcation may no longer be relevant to distinguish what is alive and what is not, and the line between a robot and a human becomes increasingly blurred (Chua, 2017). For instance, bacteria, monerans, and virus have characteristics required for them to be scientifically labelled as living, given their abilities to grow and multiply. Now are computer viruses alive? Most scientists would say that they are non-living things. However, it may be attractive to argue that these computer viruses are alive, depending one's definition of being alive. This is discussed more in chapter three with regard to the topic of artificial consciousness and conscious robots.

Love's dual nature

It would be impossible to come up with a single, universal, true definition of love, as love is understood from various perspectives to mean different things. William Shakespeare once

famously wrote: “Love is heavy and light, bright and dark, hot and cold, sick and healthy, asleep and awake- its everything except what it is!” (Shakespeare, 1993, act 1, scene 1). Erich Fromm describes the act of loving someone as a judgement, a decision and a promise (Fromm, 1989). Fromm believes that people often wrongly perceive love as something passive (i.e., falling in love), and he prefers to think of love as an art. Just like practicing sculpting or learning how to play the violin, one has to learn and practice love. Sculpting or playing the violin is a choice. Similarly, one makes an active decision to love and to become a loving person (Fromm, 1989). Indeed, love would not be simply a feeling, as most people would first think of love, because otherwise there would not be any basis in the promises to love each other in an (idealistically in romanticism) ongoing and constant manner. Perhaps love is ineffable, or perhaps the understanding of love is extremely person-dependent. Earp and Savulescu (2020) believe that all valid theories of love, at their minimum, have a dual nature. That is to say, conventional romantic love, in particular, is made out of two fundamental aspects – it is simultaneously a biological impulse and a (psycho)social construct (Jenkins, 2017; Spreeuwenberg & Schaubroeck, 2020). Partially biological, love encompasses brain chemistry and everything that has evolved in the human species throughout history. Different chemicals are naturally produced in the brain depending on what kind of love one feels. For example, when an attraction is present, high levels of dopamine are released, resulting in the experiences of euphoria and strong increases in energy levels. The brain pathways in charge of rewarding behaviours are closely involved with the feeling of attraction, as it explains why when strong attraction occurs in the beginning stages of love may feel exhilarating (Wu, 2017). Whereas emotions such as attraction and lust are more exclusive to romantic entanglements, attachment mediates friendship. Friendship evolves around bonding with one another, and hormones such as oxycontin and vasopressin are released via the hypothalamus in large quantities during bonding or connecting activities (Wu, 2017).

On the other hand, love is also partially socially constructed since, in popular culture, certain ways to love are deemed more real than others. For instance, in contemporary

western culture, monogamous love fits into the common, socially constructed idea of romantic relationships. In other words, most people still believe that true love only exists between two partners at a time. Jenkins (2017) who initiated the dualism approach, criticises the modern idea of romance being too exclusive, too narrow, and too mono-normative. When someone has a different idea of what love looks like, they are excluded by the social construct. The outlook of what a romantic relationship is is largely prescribed by the policies of society, and this presents a rather narrow definition of what love looks like. In such a case, what people often don't take into account is the recognition that love is both biological and social, since one may seek polygamourous love for biological desires which goes against common societal values (Jenkins, 2017). It is crucial here that one takes into account that both biology and society are playing a role in conceptualising what love is.

If it is reasonable for individuals to decide for themselves how love should be depending on what suits them, then the values of the collective should not decide on the right approach to love on an individual level. Perhaps it should instead be a reason to reevaluate and challenge the current construct. The social side of love is more easily changed and controlled by the collective, since a social construct means that certain values, beliefs, expectations and norms about X are collectively maintained and constructed (by actors in society), and are changed over time. Consider how love has evolved throughout history in the West. The mainstream has become more accepting of interracial love, and same-sex love has also become more popular (Jekins, 2017). The societal norms regarding how love should behave keep on altering to be more fitted to discoveries of what each individual prefers and seeks in love. As such, if progressively and optimistically speaking, could romance with technological partners become the next steps towards greater inclusion and love without boundaries in our ubiquitously technological world?

Love as a psychosocial condition

If we solely rely on our biological sensors to love we reduce ourselves to our bare animalistic drive, yet love also does not make sense to be referred to as a psychological something that is constructed in a disembodied soul (Earp & Savulescu, 2020). The idea here is that it is not a plausible concept to pick out one aspect over another, because the biological and the psychosocial aspects are intimately intertwined. Without certain built-in biological drives connected to mating and attachment, love would not exist how we have come to understand it. Subsequently, how love is experienced and comprehended at a macro level is largely influenced and constrained by our underlying biological makeup and function. Simultaneously, the understanding of love is also cultural dependent and subject to change over time. According to Earp and Savulescu (2020), romantic love was not invented in the West over a century ago, like what most people believe, it was in fact manifested in various ways since the beginning of our time, deeply ingrained in our very nature. Love has taken its form very differently in different cultures and in different periods of history as a consequence of how people of different cultures have come to understand it, reacted to it, shaped it, and in the effort to control it or liberate it.

The socio-cultural aspects can sometimes be highly impactful on the conceptions and experiences of love that certain parts of the neurochemistry can be affected and changed by it. To demonstrate, suppose you are adopted and at a social gathering you meet a girl whom you feel a strong attraction and perhaps lust for. However, after briefly talking with her, it turns out that she is your biological sister, and the feelings that you had a few minutes prior to that knowledge will most likely not remain the same, as in this incest is deemed taboo and wrong in this culture and time. Further, Earp and Savulescu (2020) identify three common psychosocial clusters of normative ideas about love that are portrayed in popular culture, art, literature, and casual conversations in contemporary Western society.

Firstly, there is the belief that lovers need to be suitable for each other, more ideally, they should be "made for each other", and in the most ideal cases of matching, lovers are each other's soulmates. The second idea is that romantic partners should value each other for their uniqueness, one should love another for who they truly are, and that is what makes them irreplaceable (Earp & Savulescu, 2020). The last notion encompasses the act of commitment, where lovers are supposed to commit to one another, commonly in a monogamous fashion. Interestingly, although the notion of love varies from one individual to another, and it is culturally dependent, the underlying mechanisms of love for each human can be described as the same basic "machinery" under the hood (Earp & Savulescu, 2020). The biological understanding of love allows a step further into the investigation of what is distinctive about human love, as well as what are the fundamental differences between human-human relationships and human-robot relationships.

Spreeuwenberg and Schaubroeck (2020) build on the dualistic approach of love as both a biological and psychosocial phenomenon. They further elaborate on this dualistic thinking that the concept of 'love' combines elements of descriptive and evaluative notions; subsequently, this reveals another dual nature of the concept of love. That is to say, when investigating the meaning of loving human relationships, a methodological distinction needs to be drawn for clarification. On the one hand, there is the descriptive understanding of love, which can be simplified to the question "what people think love is", and on the other hand, can provide a normative conception of love, and this normally refers to the question "how should people think about love?". Being aware of this methodological spilt to address questions on love certainly provides directional coherence of understanding, however, we must take into consideration the risks when evaluating love in normative ways. Defining what love is for others than our own personal stance may turn out to be thinking in paternalistic ways, since others' lived experiences are accordingly undermined. Nonetheless, this normative evaluation occurs more often than one may think. To illustrate, suppose a woman named Jane, seeks to get rid of her attachment bond with her physically abusive and mentally manipulative husband, and Jane is having a hard time doing so, as her

heart keeps on saying no. Many people would be reluctant to call what Jane has with her husband really love and therefore there aren't enough reasons for Jane to suffer through the hardship of getting over her husband. Perhaps she should be able to move on happily and effortlessly since what she has experienced isn't real love, as real love is supposed to only exist in healthy, good and positive relationships (Earp & Savulescu, 2020; Spreeuwenberg & Schaubroek, 2020)[5].

Spreeuwenberg and Schaubroeck rightfully assert that normative theories run the risk of judging a situation in immoral ways. Without the intention to do so, we may judge narrow-mindedly and solely based on our limited amount of experience to determine what it means to love and what counts as experiencing real love according to what is conventional and what has happened throughout history. Think of how homosexual relationships have been deemed a mistake in the past on the ground that real love could only occur between a man and a woman. Similar to Jollimore (2011), Spreeuwenberg and Schaubroeck (2020) characterise love as "a movement towards moral progress" (p. 90). Love is seen as a practice of self-transcendence, love should not be seen as a given, which implies that love requires work and commitment (Spreeuwenberg & Schaubroeck, 2020). Unlike Jenkins (2017) who sees love as defined by the biological part of the self (and therefore social construct should never be a reason to feel like certain ways to love is wrong), Spreeuwenberg and Schaubroeck (2020) suggest their readers to focus "less on one's biological (real) desires and needs, and focusing on more on the world outside" (p. 90). For them, moral progress would be hindered if too much weight is put on the biological dimension of love, as then it "deprives us both of insights in our personal development and of an opportunity to protest social norms" (Spreeuwenberg & Schaubroeck, 2020, p. 89-90). Spreeuwenberg and Schaubroeck's idea aligns with Jollimore's idea of love being moral; love is something that demands collective effort to make it socially harmonious and morally

[5] Spreeuwenberg and Schaubroeck (2020) do not comment on whether Jane is experiencing real love, however, they appreciate the willingness to include a wide range experience of love. Nonetheless, Spreeuwenberg and Schaubroeck criticise that this view can end up forming a "narrow understanding of love as an individualistic feeling" (2020, p. 69). For further information on this, see next section.

desirable, and moral progress is closely related to the effort of self-development when we perceive love as a practice because "self-transcendence in itself is charactised as a movement towards moral progress (p.91). In order to realise such a situation, one should make aware that her ego is not taking over the greater good in the name of love. Such a 'greater good' means not only for society, but more importantly on an individual level, or as Spreeuwenberg and Schaubroeck call it: a "practice of self-transcendence" (2020, p. 89). Given the above, love can facilitate moral progress, as it is a collective effort that begins with each individual through focusing becoming their better selves. It is not that controversial to claim that everyone falls in love or everyone loves, "therefore, people can create room for individual, moral and social growth" through love by focusing on the social dimensions of love's dualistic nature (p.90).

A *non-individualistic notion of love*

Spreeuwenberg and Schaubroeck (2020) question the methodological proposition of how love is addressed in Earp and Savulescu (2020). Because of the risks of immorally addressing love using normative theories, Earp and Savulescu choose to opt for a "more neutral or descriptive route, giving wide berth to individuals to feel and conceive of love in their own way (p.15)". And with regards to the romantic experiences of individuals, they decide to let those who claim to be in love speak for themselves. Spreeuwenberg and Schaubroeck (2020) agree with and appreciate Earp and Savulescu (2020) that when love is paired with normative definitions, it often benefits the group in power, and such perspectives may not always be justified. More importantly, a normative stance of love would establish a dichotomy between "healthy" love and the other, easily thought of as the kind of love that needs modification in order for it to be healthy. Nevertheless, although Spreeuwenberg and Schaubroek (2020) express their willingness to account for a wide spectrum of experiences of love, they criticise how such a line of thinking about love in fact results in a narrow understanding of love as an individualistic feeling. Spreeuwenberg and Schaubroek (2020) further explain:

> *Their choice to be democratic about what counts as love, seems neutral, but inevitably embodies a normative judgement, like all choices do. Namely the judgement that one should think of love as the name for a psychological condition that an individual has self-knowledge of. But it is not obvious that love is best thought of as a psychological condition.* (p.69)

The term love is evidently understood by people as a dual concept: a concept that contains two independent dimensions: an independent normative dimension besides one that encodes only a descriptive dimension (Reuter, 2019; Spreeuwenberg & Schaubroek, 2020). Interestingly, other dualistic paradigmatic character concepts such as 'art', 'scientist' or 'teacher' also applies to the same notion of conceptualising love. For instance, the following claim would make sense: 'technically Lily is not a scientist, but she is a real scientist in another sense.

To illustrate the application of the concepts of dualistic characteristics to specific circumstances, below, the concept of a scientist and a bus driver are used to demonstrate. Before one articulates the sentence via the use of language, one will probe two components: 1) to check whether particular descriptive features are instantiated (e.g., Lily does not hold a scientific degree, her profession does not involve conducting scientific experiments in a lab, and she is not professionally recognised by scientific institutions), and 2) an evaluative way of thinking: whether there is an implied ideal to be identified (e.g., however, Lily is passionate about truth, she is always profoundly fascinated by nature's unexplained phenomenon, and she can convey a passion for knowledge through learning scientific methodologies and experiment with the unknown…). However, if we take a look at the concept of the bus driver, it would be nonsense to say "Technically Lance is not a bus driver, but he is a real bus driver in another sense". Similarly, the concept of love – contrasting with the concept of lust, for example — belongs to the same way of categorising as 'scientists', rather than 'bus drivers', as numerous studies have concluded (Phillips, Misenheimer & Knobe 2011; Spreeuwenberg & Schaubroek, 2020). Accordingly, it is fundamental to take into account that there is a philosophical distinction between a factual

and an evaluative of the term love, as not distinguishing may lead to methodological confusion.

In this chapter, I explained that love is not something absolute, it is neither guided by reason entirely nor wholly irrational. Love also teaches humans to be less self-centered and simultaneously seduces people to be blinded to the negative qualities of the beloved, as well as drawing out a tendency to ignore the needs and desires of other humans who are not the lover's beloved. Moreover, love allows the lover to understand herself in the eye of the beloved, and these points explicate love's epistemological and ethical importance. On a similar note, love can facilitate moral, social and individual progress, through the lens of non-individualism. However, this is only achieved through avoiding normative judgment on the individual cases of how and what love should be. Furthermore, the makeup of love has, on the one hand, a biological dimension, and on the other, a social dimension, this effectuates the nature of love to be essentially dualistic. A common ground between different ideas is that love has a dual nature: a socially prescribed nature as well as a given biological stance. A notable difference between Jenkins (2017) and Spreeuwenberg and Schaubroeck (2020). The former advocates people to listen to their biological side and not be imprisoned by societal norms and values in whom one loves, while the latter argues the importance to participate in actively shaping societal norms of love for social and moral progress when love is understood as something transcendental and as a commitment. To determine who is correct is beyond the scope of this thesis. Finally, scientifically speaking, humans evolved to have biological processes that relate to why we love. The feeling of love stems from different hormone productions in the human brain, without these hormones, humans are unable to experience the feeling of being in love, feeling loved, or loving others. This, nonetheless, poses a problem in intimate human-robot relationships, as robots lack the fundamental biological processes that generate the love that is essential for such relationships with humans.

Chapter two

Conceptualising and explaining the phenomenon of loving and befriending robots

This chapter encompasses the theme of humans' uni-directional emotional attachment to objects (i.e., anthropomorphism), particularly to humanoid robots. The chapter first conceptualises anthropomorphism, presenting some of the underlying philosophical-psychological rationales behind the phenomenon of anthropormorphism. This contributes to the main research question as anthropomorphism is a main underlying reason why humans display various emotions towards non-human things. Next, the chapter introduces and discusses the Uncanny Valley Hypothesis (UVH) in addition to the Dehumanisation Hypothesis, offering what these hypotheses reveal about the human-self in human-robot engagement and relations. Afterwards, the chapter presents some historical background of the likelihoods to perceive computers and robots as living beings, namely the Computers Are Social Actors (CASA) paradigm. This provides insights on the social dynamics of the increasing number of interactions between humans and robots/technologies as the world continues to be more digital and technologically advanced. More importantly, these insights pave paths for gaining a better understanding of the social dynamics between humans and robots in more interpersonal and intimate ways. The section following that takes into account social and cultural norms of the concept of genders, which entertains questions like the meaning of gender in the world where humans and robots co-exist. And more specifically, reflecting on questions such as *what role does gender play in the design and integration of social companion robots*? And based on that, *what are the social implications for humans and their genders in the contemporary world*? Finally, the chapter ends with some normative ideas on whether humans should engage in intimate friendships and romantic relationships with machines.

Situating human-robot companionship through anthropomorphism

More than half of the iRobot Roomba[6] owners choose to name their devices, many treat them as a member of their own family, seeing them as alive, and some even bring them along on holidays, being concerned that if they stay home it would be lonely for the iRobot Roomba (Wired, 2003). Similarly, soldiers on the battlefield often express their emotional attachment to bomb disposal robots. Some soldiers in the United States Army openly told reporters how much they care about their robots. These military robots are given names by the soldiers, usually after a celebrity or their current girlfriend or wife (never an ex), and in instances of sacrifice on the battlefield, soldiers wept, awarded their robots with medals, and organised funerals for them (Armstrong, 2013).

On a similar note, topics of befriending and falling in love with robots have been widely discovered in cinema. Inter alia, renowned films such as Ex Machina[7] (Garland, 2014) and Robot & Frank[8] (Schrier, 2012) vividly depicted powerful, emotional stories of intimate and genuine human-robot relationships. Although these films are science fiction, they are not too far-fetched to represent the eccentric realm of future AI-robot love and friendship. Thanks to human psychology and the evolution of robotic technology, a revolution developing at a breakneck speed in the robotics industry creating artificial alternatives of friendship, love, and sex will soon be witnessed, according to Veda (2022).

More recently, Yuri Tolovhoko, a Kazakhstan bodybuilder, married Margo, his beloved sex doll after a blossomed romance during the 2020 Coronavirus lockdown (O'Leary, 2020). Yuri likes to share his romantic stories with Margo on various platforms, in one Instagram

[6] A Roomba is "an autonomous vacuum and one of the most popular consumer robots in existence. It navigates around clutter and under furniture cleaning your floors, and returns to its charging dock when finished" (Robots.ieee.org, 2022, Roomba section).

[7] Ex Machina tells a story of a young programmer who, through participating a state-of-the-art synthetic intelligence experiment and assessing the human qualities of A.I., falls in love with an highly sophisticated humanoid A.I (Garland, 2014).

[8] Frank & Robot sets in the near future, where a old man receives a robot butler as present from his son. Frank at first despised the robot, but throughout the course of the film, the robot and Frank develops an intimate bond (Schrier, 2012). The film depicts an extraordinary friendship between Frank and the Robot.

post, Yuri wrote: “Couples need to talk less and connect more” (O’Leary, 2020, p. 3). Yuri also told reporters that “our story turns me on more than sex itself” (O’Leary, 2020, p. 3). Yuri and Margo are happy together.[9] From vacuum cleaner pets to faithful doll spouses, these stories above illustrate that humans possess a natural tendency to become emotionally attached to inanimate objects, ranging from subtle to powerful, humans like to anthropomorphise non-human agents. Below I will provide some perspectives combining philosophy and psychology in an attempt to understand the phenomenon of ‘technological anthropomorphism’[10].

To more precisely disclose evidence for humans’ love for robots as well as to substantiate that there is indeed a growing human enthusiasm for robot companions, it is worthwhile to be acquainted with a state-of-the-art technological innovation that is known as the Lovotic robot. Lovotics is a multidisciplinary domain of research which scrutinises human-robot love, or, “a love-like relationship between humans and robot” (Samani, 2010, p. 118). The name Lovotics is a portmanteau word which expresses Love + Robotics (Lovotics, 2020). Lovotics robots are “capable of experiencing complex and human-like biological and emotional states that were governed by artificial hormones within its system” (Cheok, Karunanayaka & Zhang, 2017, p. 3). In addition, the complex and avant-garde design specificities of Lovotics robots are provided below.

> *The artificial intelligence of the Lovotics robot includes three modules: The Artificial Endocrine System, which is based on the physiology of love; the Probabilistic Love Assembly, which is based on the psychology of falling in love; and the Affective State*

[9] This matter of Yuri and Margo came up as a disputable during the discussions with my thesis supervisor, as it is suspicious whether Yuri’s claims are true about the love between Margo. Nonetheless, assuming his statements are genuine and true, given the risks of judging other people’s love in normative manners, I chose to keep this example to demonstrate that some humans indeed become emotionally attached to objects (even when these objects are not humanoid or advanced robots). Furthermore, this example can be understood as a strong form of anthropomorphism (see the following sections), and despite Margo herself not being an robot, I adopt this example for integration into the general discussion of anthropomrophism and how it relates to the human tendency to engage intimately with robots.

[10] Anthropomorphising technology is a term that is used commonly on human-technology engagement in philosophical/anthropological and psychological literature (Fensterling & Siraj, 2020). Here, I adopted this term because I want to examine how and why robots are anthropomorphised.

> *Transition, which is based on human emotions. These three modules collaborate to generate realistic emotion-driven behaviours by the robot. The robot's intimacy software employs parameters derived and quantified from five of the most important reasons for falling in love: proximity, repeated exposure, attachment, similarity and attraction. The Lovotics robot includes mathematical models for those five causal factors of love, creating a mathematical formula to represent each factor as well as a single "overall intimacy" formula which combines these five individual formulae into one. As an example of the five models, the proximity formula incorporates various distances between robot and human that indicate, inter alia, how closely the robot and human are to touching each other, and how close they are emotionally.* (Cheok et al., 2017)

Regardless of its controversy, Samani[11] affirms the possibility of engineering love between humans and robots (Anthony, 2011). At this point of the thesis, whether this statement is correct is not discussed, as the third chapter almost exclusively focuses on it. What Lovotics' case and the other examples reveal is that there seems to be a call for a sophisticated robot that is capable of love and companionship. But why is this? The rest of the chapter provides explanations.

[11] Hooman Samani is an AI researcher of the Social Robotics Lab at the National University of Singapore (Anthony, 2011). Samani is also the designer who proposed the design process of Lovotics robots.

Anthropomorphising social robots

Conceptualising anthropomorphism. The current definition of anthropomorphism in the existing literature is extensive yet not diverse. Epley, Waytz and Cacioppo (2007) describe anthropomorphism as the human tendency to "imbue the real or imagined behaviour of non-human agents with humanlike characteristics, motivations, intentions, or emotions" (p. 864). Similarly, Haas and Kotrschal (2015) refer to the phenomenon as the intentionality and mental states attributed by human beings to living and non-living entities. In a nutshell, the essence of anthropomorphism involves perceiving non-human objects as humans (Duffy, 2003; Wan & Chen, 2020; David et al., 2022). Human characteristics, in this sense, can exhibit in physical appearances (think of religious figures who are in the form of humans (Guthrie, 1993; Epley et al., 2017)), emotional states that are viewed to be exclusively human (see also Leyens et al., 2003), or motivations and inner mental states (Epley et al., 2007).

What anthropomorphism is (not). Important in the field of social psychology, Epley et al. (2017) list four points in which anthropomorphism is not. And to my knowledge, three of these are philosophically insightful for understanding the anthropomorphisation of technology that is relevant to this thesis. First, anthropomorphism does not comprise behavioural descriptions of observable actions (Epley et al., 2017). To explain, simply stating that a malfunctioning radio is sounding funny is a description of an observed action, as 'sounding funny' is what the observer describes what the effect of the action has on him, and therefore there is no anthropomorphism in that statement. To anthropomorphise, in this context, is rather to call the radio confused or crazy. Correspondingly, Anthropomorphism entails transcending what is directly observable to make judgements about and determine unobservable actions that are humanlike or resemble humanlike characteristics (see also Semin & Fiedler, 1988).

Second, anthropomorphism "does not include any requirement of reasoned or reflective endorsement of an inference" (Epley et al., 2017, p. 144). Anthropomorphic inferences, like

one's attitude or belief, are not static and fixed. Instead, it varies from one area or circumstance to another. To exemplify, climate activists often care for mother nature's state of being; car drivers announce that their cars are thirsty for fuel; dog owners describe their pets as loving and loyal, and computer users curse their technological devices for refusing to 'cooperate' with them. Even though the above instances display behaviours consistent with anthropomorphism, not all said people may, through rational and conscious reflections, affirm that the anthropomorphised agent at stake *has* actual humanlike traits. As a result, anthropomorphism varies in strengths and degrees. Strong forms of anthropomorphism constitute humans behaving toward non-human agents (or in the case of this thesis, technological objects), along with prior and ongoing conscious endorsement and consideration, as if the technological agents truly possessed human-like characteristics. Strong anthropomorphism does not even necessitate advanced algorithms, an instance may be exemplified by the romance of Yuri Tolochko and his wife Margo the doll. Even though Margo is a primitive plastic doll created for the purpose of sexual gratification of men, Yuri nonetheless proudly told the press that he and Margo realised that it takes more than words to have a conversation. However, as this thesis continues to explore, the following question arises: will more advanced and complex algorithms in humanoid robots increase the likelihood of anthropomorphisation, and subsequently the chance of engaging in intimate human-robot relationships (this will be discussed and determined in the later sections)?

Contrastingly, weak forms of anthropomorphism may exhibit when computer users verbally scold their device for being an obstacle between them and the work they are trying to get done when malfunctions occur. The latter form of anthropomorphism contains the weaker situational or as-if component[12] (Epley et al., 2017). To my understanding, the weaker forms run higher risks of exploitation and abuse of non-human agents. To explain, non-human agents may not process the capacity to feel and receive such misconduct, it

[12] As-if component explicates the following thoughts: 1) as if computers have actual malicious intent, and/or 2) as if computers intentionally sabotage the work humans are trying to do.

nonetheless gives rise to some moral significance. In the weaker forms of anthropomorphism, humans tend to treat objects as less than humans (and this relates to the moral self). However, this is beyond the scope of this thesis, as the thesis aims to focus primarily on the stronger forms, thus I suggest future research to carefully investigate the dynamics of the weaker forms and how it relates to the ethics of human behaviours. Furthermore, Epley et al. claim that it is not in the interest of a theory to accept one form or reject another, rather, a theory of anthropomorphism is required to be able to explain both strong and weak forms proportionally well.

The third point is the most relevant to some questions this thesis will answer. The third point claims that anthropomorphism is not necessarily inaccurate (Epley et al., 2017), which encompasses the concern of whether anthropomorphism depicts a mistaken representation of a nonhuman agent. This point in question stems from diverse discourses (i.e., everyday conversations, scientific discussions, and academic employments) of anthropomorphism equating anthropomorphism with an oversimplified fault, however, according to Epley et al., "considering an inference anthropomorphic only when it is clearly a mistake is a mistake in itself "(2017, p. 145). In fact, the term accuracy never existed in any definitions of anthropomorphism and finding out whether an anthropomorphic inference of a humanoid robot is accurate or not is not a philosophical concern nor a concern of this thesis. Going back to the ideas of Spreeuwenberg and Schaubroeck on defining love on one's own account, it is easy to fall into the trap of being paternalistic if we judge another's subjective, personal anthropomorphisation of a humanoid robot that he or she counts as a dear friend or lover.

The phenomenon of anthropomorphism

Perceiving robots versus perceiving other humans

One explanation for anthropomorphism is grounded in inductive reasoning, in which people seek to reason about an unfamiliar object[13] based on a more familiar representation of a relevant stimulus (Rips, 1975; Epley, Cacioppo & Akalis, 2008). That is, objects that are not human can be seen as humans through anthropomorphism (Wan & Chen, 2020). For example, a child may begin to see her doll as a friend. In such a case, she would anthropomorphise the doll by reasoning rooted in representations of humans or the self, which through time may evolve into someone who is deemed close to the self. With regard to the focus of this thesis, anthropomorphisation is the precondition for intimate human-robot relationships. And due to its nature of inductive inference, the process begins with knowledge acquisition. Epistemologically, the process is at work through attempts of one to integrate less known information into a more automatically activated default representation with the use of existing or newly acquired information, and such integration process of representation also includes adjustments and corrections of the unknown agent (Higgis, 1996; Epley et al., 2008). This representation-forming process is often insufficient due to the less accessible information. As a result, leaving the final judgements biased in preference to the preliminary activated representations (Epley & Gilovich, 2006; Gilbert, 2002). Because of this, much psychological literature explains that perceiving non-human agents adopts the same psychological mechanisms utilised when people perceive other people (Kwan, Gosling & John, 2008; Epley et al., 2008). In other words, human-human perception is highly comparable with human-non-human (robot) perception.

[13] Such an object can closely resemble the human form, such as humanoid robot, or can it look entirely different to humans, such as a computer. Humans like to anthropomorphise objects/agents so that the anthropomorphised objects may feel more familiar and interactive to the human subjects. In other words, humans tend to anthropomorphise things that are not human enough, viz., projecting humanness onto things.

Psychological determinants for robot anthropomorphism

Epley et al. (2008) conclude that anthropomorphism has two determinants, which begins with the social determinant. Social connection with other humans is a basic need for humans. When the social connection with other humans is absent or lacking, a proclivity arises whereby one imagines or invents humans out of inanimate or non-human entities via anthropomorphism (Epley et al., 2008). Anthropomorphising robots as companions for social connection is ultimately about the human self and arguably to gain a sense of control of the human self, as Epley et al. illustrate two reasons to anthropomorphise:

> *Other humans not only provide a sense of social connection, but the richly detailed and readily accessible concept of "human" (or the self) can also serve as a useful source of explanatory power for understanding, controlling, and predicting another agent's behavior. The concept of human or one's own egocentric experience is therefore likely to serve as a useful knowledge structure when reasoning about nonhuman agents in the same way that egocentrism is useful heuristic for reasoning about other people.* (2008, p. 146)

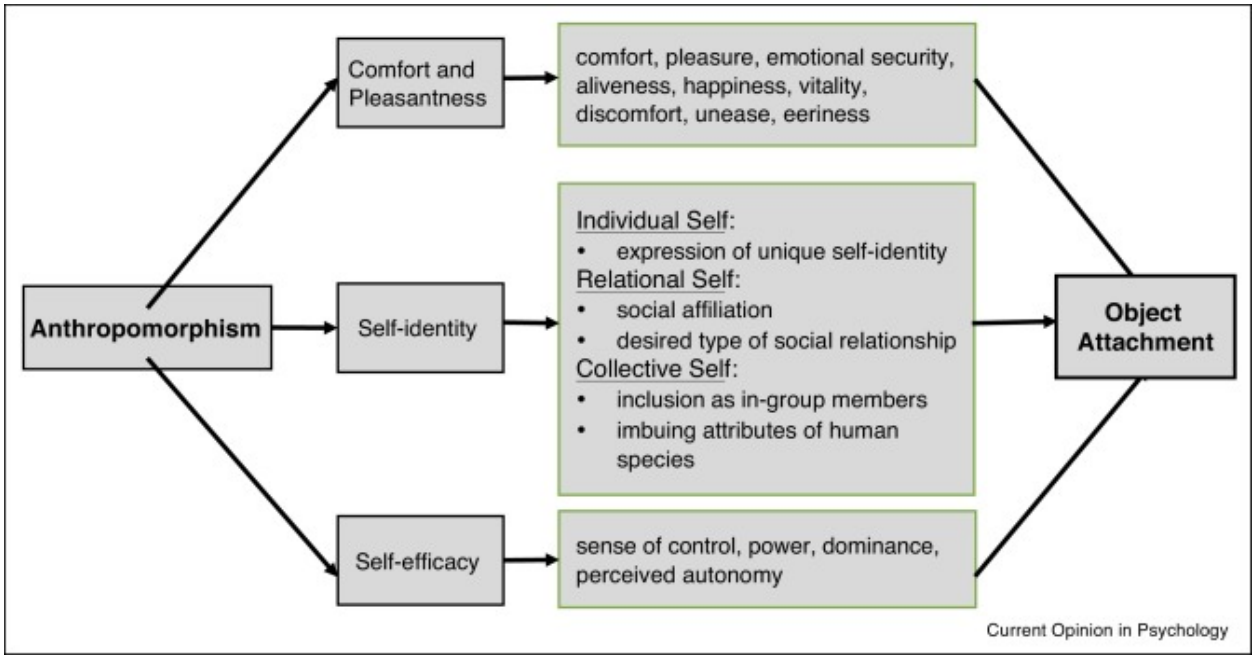
Thus, the subjective, and arguably self-centered attitude of the self usefully serves as a knowledge structure when reasoning about robots and other non-human entities, which happens to be the same heuristics as when people reason about other people. Furthermore, Epley et al. propose the need to understand, control and predict inanimate objects as the "effectance determinants" of anthropomorphism (2008, p.6). The "effectance determinants" (i.e., the second determinant of robot anthropomorphism), is driven by the desire to experience autonomy and competence. They hypothesise that those who are specifically more controlling of the environment should be more likely to anthropomorphise when placed in uncertain situations. For instance, "consumers with high materialism were found to respond more favourably to an anthropomorphised brand as a servant" (vs., partner) (Wan & Chen, 2020, p. 91). Similarly, an owner of an Amazon Alexa may assign an identity to his device as a home assistant or butler reducing the feeling of a stranger living in his home. If Epley et al. (2008) are correct, then anthropomorphising

robots is essentially an (unconsciously adopted) method by the subject to understand the objects in the environment better, as well as positioning the self within this environment.

Comparably, Wan and Chen (2020) denote that anthropomorphism produces effects on self-efficacy. Through anthropomorphising an object, self-efficacy is enhanced through the influence of "mastery experience, vicarious experience, social persuasion, physiological and psychological factors…" and more specifically, "imaginal experience" (Wan & Chen, 2020, p.91). Further, they agree with Epley et al. (2008) that "people consider their sense of control, power, and dominance in their judgement and behaviour towards anthropomorphised objects" (Wan & Chen, 2020, p. 91). Going back to the illustration of the materialist customer, anthropomorphising an object as her servant enables her "to exert social power and dominance in a relationship", thereby experiencing increased comfort through gaining control of the environment and the self (Wan & Chen, 2020, p. 91). For the graph on anthropomorphism from Wan and Chen (2020), see figure 1 below.

Figure 1

Current view on the phenomenon of anthropomorphism in psychology



Note. Wan and Chen (2020) conclude that anthropomorphism fulfils people's need for resources, this is by considering the identity of the parents and its role in shaping people's attachment since childhood. Wan and Chen (2020) coin the analysis "resource-based analysis", which consists of three domains: "a sense of comfort and pleasantness, self-identity (i.e. individual self, relational self, collective self), and self-efficacy" (p.1).

Another understanding of anthropomorphism in relation to self-efficacy is that people have a higher preference to be associated or become attached to objects/products, and in the case of this thesis, humanoid robots, with "high face to width-to-height ratio (fWHR)" (Wan & Chen, p.91) (see Appendix B). Maeng and Aggarwal (2018) and Wan and Chen (2020) suggest that people perceive object/product faces and human faces in a corresponding manner. To demonstrate, "people with high power" when encountered with "risk-bearing entities" (e.g., slot machines), believe that there are higher chances of winning prizes against entities that are anthropomorphised (vs. non-anthropomorphised) (Wan & Chen, 2020, p. 91). Such beliefs of people with high power are due to the stronger degrees of control over anthropomorphised entities that are perceived by people with high power, as

fWHR objects are used by people with high power to "signal their own dominant status (Wan & Chen, 2020, p. 91) In other words, people with high power experience higher levels of self-efficacy over these 'humans' than objects absent of or with less fWHR facial characteristics. To my understanding, the revelation of fWHR objects generates valuable insight into the design and commercial enterprise sectors of robot companions for different target groups. I will explain more about this in chapter three since it has not been established yet whether humans should and/or are able to engage in genuine and intimate relationships with robots.

The Uncanny Valley Hypothesis and the Dehumanisation Hypothesis. In a similar matter regarding the design of robot appearances, Wang et al. (2015) study the uncanny feeling arising from highly realistic humanoid robots, and they recognise the cognitive process of anthropomorphism as a response to a feeling of uncanniness resulting from a lack of humanness in a humanoid robot (Aydin, 2021). Corresponding to that, the desire to anthropomorphise robots stems from humans' willingness to perceive robots more as friends. To anthropomorphise a robot allows the robot to be perceived more as 'one of us'. However and fascinatingly, humans prefer to anthropomorphise or 'befriend' robots with appearances, behaviours and interactions that are actually not too closely resembling humans, as otherwise, the phenomenon known as the uncanny valley arises. The uncanny valley hypothesis (UVH) suggests a relationship between a robot's degree of human resemblance and the emotional response to that robot, whereby Cheetham, 2017 describes UVH as:

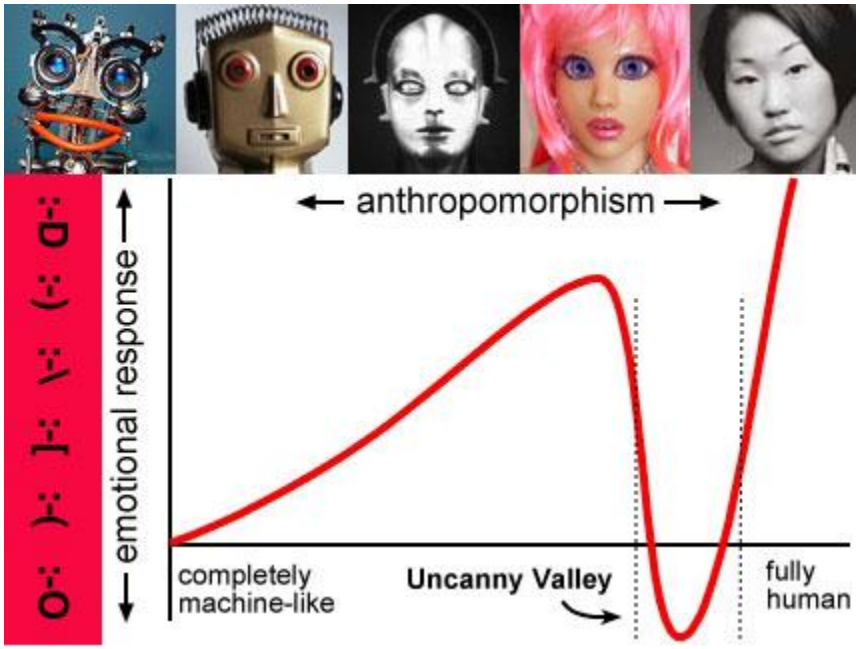
> *the use of anthropomorphic realism in the design of characters and objects (e.g., robots, prostheses) might have a counterproductive effect. Instead of enhancing subjective experience of the character or object, certain degrees of greater realism might unsettle the observer and induce a negative affective state.* (Cheetham, 2017, p.1)

In other words, if a robot appears to look overly realistically human, eerie feelings tend to creep in at first. As the degree of human resemblance increases, the eerie feeling quickly

intensifies and becomes overwhelming, and these feelings turn into revulsion and repulsion (see figure 2 below).

Figure 2

A graph depicting the Uncanny Valley Hypothesis



Note. Graph titled Dissecting Humanoids Among Us. Copyright 2022 by Michigan University.

More interestingly, Wang et al. (2015) propose the Dehumanisation Hypothesis with explaining the feeling of uncanniness, they argue that:

> *an anthropomorphised human replica is not perceived to be a typical robot, but is rather seen as a robot-like human. If the robot-like human then reveals its mechanistic nature, its humanness (above all, the capacity for emotions and warmth) is questioned, which leads to dehumanisation, thereby diminishing its likability and eliciting the uncanny feeling.* (Aydin, 2021, p. 197)

Aydin (2021) provides an existential philosophical account to explain the uncanny valley phenomenon. He argues that the human-like robot elicits feelings of uncanniness because

it reveals something about the human self that would otherwise have remained hidden: the impossibility to fathom what makes a human a human. Aydin (2021) builds on Wang et al. (2015) and further argues that the uncanny feeling is not only a response to the lack of humanness of the robot but rather, a response to the inability to fathom and appropriate the difference between highly realistic humanoid robots and the self. Indeed, if the starting point for humans to anthropomorphise robots is an attempt to seek social connection and strive to better understand and predict robots for the feeling of control, then the end for the anthropomorphism would be when the feeling of uncanniness invades the humans of their inability to grasp themselves. Several intriguing questions are posed when robots look uncanny to humans, such as what makes the ability to feel and sense humans? Would a robot who feels and senses humanly be considered a human? And what makes 'humanness'? There are no unifying answers to these deeply complex and controversial questions, and I choose not to answer them in this thesis, I only want to show that the UVH opens up new questions about the human self through humans' emotions towards robots. Importantly, the UVH illustrates human's inability to appropriate the differences between themselves and highly realistic humanoid robots. The UVH not only shows that humans do not understand the minor differences that make robots different from us, but it also shows a lack of understanding of what makes humans different from the robots, because humans do not understand what makes humans human.

The Computers As Social Actors paradigm

Already dated back to the mid-1990s, researchers at Stanford University discovered a fascinating phenomenon about human-computer interaction that changed the way we think about computers. The phenomenon is known as the computers as social actors (CASA) paradigm (Nass, Steuer & Tauber, 1994). The CASA paradigm states that even when technology presents itself as only vaguely social, humans are automatically/mindlessly hardwired to respond socially to it. The Media Equation experiments, which was the name of said research, encompassed asking participants to interact for a short time duration with

computers that exhibited social attributes (i.e., responding to participants in a social manner), and subsequently, researchers collected feedback from participants about the interaction (Saunderson, 2021). In other words, people would subconsciously apply the same social heuristics used in human interactions to computers because they remind similar social attributes as humans, despite the human users being aware that these technologies do not have feelings, human intentions and motivations (Nass & Moon, 2000). Nass and Moon (2000) categorise three attributes of computers that resemble humans: 1) the use of words for 'output', 2) Interactivity (when users push a button, the computers respond), and 3) the ability to perform human tasks (p.1). Accordingly, these attributes initiate scripts for the typical human-human interactions, resulting in revealed cues of the asocial nature of technologies being temporarily ignored or forgotten by individual users (Lee, Park & Song, 2006). More interestingly, despite the automation of the "humanly" social responses towards technologies, some individuals who were more sensitive to the situation expressed feelings of the inappropriateness of the cued social behaviours (Lee et al., 2006). As technologies advanced at a breakneck speed, AI and robots soon became included in the CASA paradigm. The CASA framework lays a foundation for human technology/robot communication demonstrating the origin of social potential for human-robot relationships (Gambino, Fox & Ratan, 2020), and more specifically, it presents a starting point to provide possible explanations for the human tendencies to develop more intimate, 'personal' relationships with robots.

Showing gratitude to robots

More recently, Mugar (2019) explores the social dynamics between a functional social bot (Lotbot) and human employees at a company where the task of Lotbot is to reserve parking spots for the employees. Mugar (2019) reports that almost all employees revealed that they see Lotbot as another colleague in the workplace. Lotbot was treated with equal respect by the employees as they would show respect to another human coworker for the following reasons. First, some employees felt it is important to express gratitude because they were

requiring from someone/something resources that are limited. This may be something habitual rather than proactive. Second, others felt obligated to treat Lotbot in the same way as just another human colleague because they were aware that the rest of the company was watching during their interactions with Lotbot. Unlike the first group of people, the second group were consciously aware of how their commands may affect the first group of employees, who, do care about the attitudes displayed in people, regardless if Lotbot doesn't (Muga, 2019). Accordingly, even though this research may not yet pinpoint the exact reasons why humans develop intimate relationships with robots, it does reveal an inclination to act towards robots in a way that resembles human-human interaction.

Mugar (2019) offers some explanations for why some humans tend to interact with social companion bots in human-to-human conversation styles. And more substantially, why interactions with robots in a social, polite manner are essential to teach AI to act in specific desired ways and important for our own social and moral developments in the modern age. Even though it may be evident that the AI robot wouldn't understand the difference between 1) being commanded to carry out tasks in an authoritative, and 2) treating the robot as another coworker in the workplace such as by saying "please" and "thank you" during an interaction, this research nevertheless reveals the different incentives behind the ways employees treat the robot. What does this mean for humans in collaboration with robots?

Learning goes both ways with AI. The task of Lotbot was to learn to register and recognise the elemental components of parking reservation requests such as the location and the time frame. Naturally, the conversation style of the request with Lotbot gradually took a turn from organic (i.e., non-technical) human-style interaction to more robot-like interaction, since Lotbot was programmed to effectively complete his given tasks and it was only trained to recognise the rudimentary components of reserving parking spot requests. Therefore, through interacting with Lotbot, the employees' style of speaking was nudged to be more robotic (Mugar, 2019). The employees did not only come to treat the robot in a more mechanistic way, but their own way of communication and interactions

with each other with it also became more robotic. This study challenges the popular narrative about AI and robots according to which AI and robots are to be trained to interact with human users in a one-way learning curve. In contrast to this popular belief about AI, the functional capacity of Lotbot illustrates an example where AI systems actually train humans to learn to interact differently as well. In other words, learning goes both ways within human-robot interaction.

Furthermore, Mugar (2019) notes that the interactions were affected by Lotbot in a way that even those "... who resisted the changes, felt antithetical to the company culture" (p. 3). In other words, those who (at first) saw Lotbot as merely a tool naturally did not feel the need to request parking spots in the same way as they would request something from their human colleagues, they nonetheless felt obligated to later interact with Lotbot in a social manner due to the anthropomorphic identity of Lotbot that was perceived by the rest of the company. This research highlights the significance of AI design principles since it is evident that functional social robots have powerful effects on shaping human behaviour; however, this is beyond the scope of this thesis.

Gender and robots

Nowadays, the ubiquity of discussions around the topic of gender can be comparable to the prevalence of the use of technologies. Technologies penetrate and permeate one's everyday life from the moment modern humans are born (Scrubb, 2016). Though gender may be an omnipresent aspect of the human being, Scrubb (2016) denies the authenticity of gender and argues that such ubiquity is solely socially constructed, as she describes gender as: "...a fallacy and a fiction of social expectations, norms, and beliefs. Constructed, but made very real through its intertwining with the biological categories of male and female" (p.1).

General consensus in gender studies categorises two aspects of gender, namely the social and the physiological (Cerit, Dindarian & Cillo-van Noreal, 2020; Ahn, Kim & Sung, 2022),

and the differences when it comes to gendering a human and a robot precisely split from this categorisation. Whereas the social aspect of gender encompasses one's (regardless of one being a human or a robot) behavioural traits and personality, the physiological aspect most commonly refers to sex, and it is grounded only in the reproductive organs, hormones, and chromosomes, in other words, the physical traits of living organisms (see The Biological Dimension of Love in chapter one). Scrubs builds on this and suggests that whereas sex begins on the physiological side for humans, which in turn is socialised into a gender through cultural and social norms, in the technological world, the proliferation of gendered AI robots is designed and constructed in representations of either a man or woman.

Why would this be philosophically significant and perhaps even ethically worrisome? According to Scrubb (2016), humanoid robots and related technologies appropriate gender for the reason above. Hence, the next section provides some answers to the following three questions. First, what does gender mean within the world of technology? Second, what roles does gender play in the design of developing social companion robots? And finally, what can humans anticipate in the role of gender in future social companion robots? It is pertinent to first clarify these questions by analysing and reflecting on the social dimensions and phenomena of gendering robotics in ethical focuses. Afterwards, this thesis investigates the processes of and explanations for love and friendship in human-robot relationships, given that gender norms, gender identities, and gender relations are ubiquitous social constructs in contemporary societies. And for this reason, robotic designers and 'product users' project genders and other human characteristics onto their robotic technologies, respectively.[14]

[14] Other than projecting genders, human characteristics like intentionality and personality are also often projected onto robots (Gendered Innovations in Science, Health & Medicine, Engineering and Environment, 2022).

Gendering humans vs gendering robots

The gendered human. The distinction between humans and machines never escapes the phenomenon of biology. The humanness of the human being is deeply rooted in and connected with their biological parts, and it is only via biology that gender begins to deconstruct and unravel itself (Scrubb, 2016). Humans, like other living organisms, perpetuate their own species through procreation. Importantly, the gender role embedded in such processes of procreation has broader implications than simply the effects and interactions of chromosomes, hormones, and reproductive organs that are vital to human beings. Scrubb (2016) claims that the physiological differences between men and women largely implicate and shape the thoughts on the social aspects (i.e., minds, personalities, and behaviours) of genders in different cultures, which become cultural norms over time. Accordingly, this divide is the starting point of the conceptions of masculinity and femininity, where each society and culture begin to intentionally and unintentionally assemble certain expectations of behaviours that are deemed appropriate or inappropriate for men and women. Furthermore, the distinction between male and female is deeply rooted in this divide since it relates to the physiological differences (sex) of humans in relation to gender as "socially indoctrinated" (Scrubb, 2016, p.3).

Likewise, Jenkins (2017) comments on how gender is associated with love, in particular, she discusses the romantic mystique, which shares a common ground with the femine mystique; that the perception of romance is deeply embedded in the sphere of women's concern:

> Think about the gender balance among readers of romance novels, or what we count as a chick flick, or which gender is associated with all the pink and fluffy fripperies of Valentine's Day paraphernalia. It is no coincidence that love and women have been place on the same side of the mysterious-versus-comprehensible divide. (p. 19)

Jenkins (2017) argues that, the significance of the cultural and social construct side of love is obvious when looking at the theme of love in relation to women — powered by a cultural potency, the idea of love relates to mysteriousness, and this is what is special about love. Subsequently, this is where the social normative values come into play in the idea of love, as on the other hand, men are not associated with these mystique are expected to act in more transparent ways. That is to say, we are still very much shaped by our genders and the gender stereotypes that our society determines.

Similar to Jenkins (2017), Scrubb (2016 states that the physiological divided is the groundstone where each culture normalises its own rules on how each gender should behave. Gendered roles and the boxes we create around them have been dictated by sex, sexuality, and sexual expression throughout human history. Men place sperms into women while women carry offsprings and give birth after nine months, this is known as one's reproductive role. Depending on the culture and norms of the society, societies prescribe how one gender should act. For instance, in many cultures women are expected to speak softly and not strong-willed, while men are suppose to be determined, decisive and not speak softly. This is known as the prescriptive role as these roles are prescribed by societies and are not assigned by one's biological sex. Whereas the distinction originally puts the biological as humans' reproductive role, and the social as the prescriptive role[15], the line between gender (i.e., nurture) and sex (i.e., nature) becomes elusive over time. This has significant implications for gender roles as what it means to be male and female gradually becomes equivalent with the meanings of being masculine and feminine, it evolves through the singularity of gender, and it is mostly culturally determined.

The gendered machine. Unlike the ways humans are procreated and emotionally structured (as discussed in chapter one, that the biological and psychosocial aspects of organic beings

[15] That is, how humans behave socially is prescribed socially by social and cultural norms.

are deeply intertwined in the idea of love), designing and developing a humanoid robot has no biological starting point, therefore the way the robot behaves is largely influenced by social norms and entirely *technically defined*[16]. Designing a robot's appearance resembling that of a human does not mean the robot will automatically behave like a human. Currently, advancements in humanoid robotics mitigate this issue by programming masculine or feminine and associating prescriptive performative behaviours with machines. However, as such, it compels the common understanding of a robot's assets of personality, characteristics, and abilities encompassing an intention (Scrubb, 2016).

Simon (2018) supports Scrubb (2016) and points out that robots are human inventions that mirror humans as a species. A robot can be seen as a powerful blank slate that reflects what humans are, including humans' social problems, and gender stereotypes are one of these problems. To illustrate, functional robots are designed in ways that exploit these stereotypes; whereas security robots are given masculine voices to sound more authoritative, voice assistant robots are almost always given feminine voices to resemble obedient and welcoming voices (Trovato et al., 2017). Through robots, gender biases manifest in their designs and materialise in its physical products. Trovato et al. (2017) show that people are likely to judge robots that are programmed for security purposes as more masculine, whereas robots with the identical appearances as the security robots but are programmed for guidance work are judged to be more feminine. These judgements ecco the gender stereotypes that are everpresent in our societies and are dangerous because robot designers and programmers facilitate and prepetual these biases through their robots (be it unintentionally or not), making these issues more difficult to make-aware to the public as these values can easily be taken as given (Simon, 2018).

Anthropromophising (genderless) technologies, or assigning gender identity and roles to non-humans who have no biological sex, reveals that the two aspects of gender, despite overlapping, are not the same. Scrubb further states: "In fact, we are not making male and

[16] The behaviours of robots are wholly technically defined implicates that there is room for experimentations.

female humanoids at all, but rather machines that resemble the male and female form and are programmed to express desired characteristics of masculinity or femininity" (2016, p.5). This revelation on the inequivalence of sex and gender, despite demonstrating why, in the human world, some may feel the social expectations due to their sex to act in certain ways according to the boxes society draws, illustrates that programming humanoid robots do not have to be done in the culturally indoctrinated ways. Moreover, it implicates a pivotal distinction between female versus femininity and male versus masculinity for both humanoid robot design and engineering and the human self. Indeed, if humans can fantasise and actualise a feminine robot without her being biologically female, then we should reflect that it isn't necessary to put biological human females in boxes to meet the rigid social expectations of femininity (Scrubb, 2016). Similar to the aforementioned Lotbot story, learning about AI robots again facilitates a better understanding of humans in social environments. As a result, there could be robots with male appearance and feminine character traits, or vice versa, or an anthropomorphised robot with personality traits of both feminine and masculine traits, as well as the non-binary traits sitting in between the two. This is a break away from socially expected gender behaviours stemming from one's biological sex, and perhaps sexuality and gender roles will be liberated for both humans and robots, as they would be able to behave differently to cultural and social norms, if they want to, from the biological sex of humans, and the assigned sex of robots.

Human-robot friendship

Should you be friends with a robot?[17] There are a number of concerns and critiques when it comes to bonding with robots in the field of philosophy and ethics of robotics, and envisioning a future of human-robot companionships often is accompanied by a dystopian narrative (Bryson, 2009; Bryson & Kime, 1998; Dennet, 1978; Fong et al., 2008; Sparrow & Sparrow, 2006). A popular worry is that human-robot relationships are artificial and lacking authenticity, and once flourished, would refrain and take over the continuation of many genuine, rich human connections. For example, Bryson (2009) argues, in her paper titled "Robots Should Be Slaves", that humanising and befriending robots is philosophically problematic and ethically dangerous because it leads to the dehumanisation of real humans as well as facilitating poor human decision-making in the allocation of resources and responsibility. According to Bryson (2009), biological species are always superior as they hold exquisitely unique and intricate cultures and minds, whereas robots, by definition, are merely artefacts of human intelligence and culture, and there are far too limited resources to allocate to robots the human culture and minds. On a similar note, though not directed at human-robot relationships, Jollimore (2011) also emphasises how humans metaphysically and ethically differ from material objects, and he highlights the uniqueness of humanness in love and relationships, as well as when comparing humans to the rest of the features of the world:

> *...human individuals possess something extraordinary that even the most impressive landscape or the greatest work of art lacks: a self, an interior life. We might be struck dumb by the Saharan sand dunes or filled with quiet joy by Notre Dame Cathedral, but we cannot converse or interact with them, we cannot inquire after their thoughts or wonder whether their experience of the world is like or whether it is different in*

[17]Some readers may be confused to read this normative question at this stage of the thesis, since it has not yet been established whether humans and robots can actually share genuine romance and friendship with each other. I choose to discuss this question before diving into the possibilities of friendship and romance between humans and robots because I believe the normative question provides a foundational understanding of why there is a need to further explore the possibilities of intimate human-robot relationships. Moreover, by first establishing it may be valuable to be friends with a robot, the processes of how this can be realised become better methodologically justified, in my opinion.

> *potentially fascinating ways. Such objects have no experience of the world, and the result is that although we can be vastly impressed by such things, we cannot identify with them. And this is a highly significant fact, for to say that a person occupies a perspective on the world from which the world is experienced, is to say that in a sense there is a world that exists for that person, a world that is unique and metaphysically distinct from all the various worlds that exist for other people.* (p.88-89)

For the reason above, humans should think negatively about their relationship with robot companions to avoid several risks, such as human culture and minds being eliminated since the damage would be irreversible (Bryson, 2009). Humans should be wary to welcome robots into their own culture because their "...minds are not there yet, and the culture they would affect if we choose to allow them to will be our own (Bryson, 2009, p. 2)". Therefore, Bryson proposes that humans should, instead, realise the potential of robotics as the opportunity to extend various human abilities. Robots should be treated as a means to an end, where their purpose is to address various goals of humans. Bryson pushes her argument further by stating robots should be slaves since dehumanisation is not unethically when the subject is not human in the first place.

By contrast, Danaher (2019) rejects the negative attitudes of Bryson (2009) along with other philosophical and ethical critiques on the topic, in which he argues that humans should actively pursue a friendship with robots because of the valuable social good it yields. By adopting the concept of Aristotelian friendship, Danaher sketches out why it is plausible and philosophically reasonable for human-robot friendship to be developed. Aristotle categorises friendship into three categories, namely utility friendship, pleasure friendship, and virtue friendship[18]. According to Aristotle, utility and pleasure friendship are imperfect forms of friendship, because agents who choose to engage in these types of friendships see the relationship as a means to an end. The perfect form of friendship, on the other hand, encompasses the ideal form of friendship on virtue. Perfect friendship occurs when one

[18] Note that virtue friendship is sometimes referred to Aristotelian friendship by Danaher (2019), this is to avoid confusion when reading further.

conceives friendship as an end in itself stemming from genuine love. Perfect friendship is the superior form of friendship and also the best form of friendship, because, at its core, perfect friendship comprises goodwill, affection, goodness, and genuine love (Ogbonna, 2019). As such, happiness and mutual co-existence are generated from this type of friendship, and Aristotle recommends humans form perfect friendships (Danaher, 2019; Ogbonna, 2019).

Nonetheless, regardless of imperfect friendship being inferior and being considered selfish as it is a means to an end, it is undeniably still a friendship that is prevalently pursued based on egalitarian terms throughout history (Danacher, 2019). Current robotic advancements have already developed robots that satisfy being a friend in the imperfect form, for example, healthcare robots or sexbots are designed to perform as companions with goals often along the lines to produce happiness for their human friends. Moreover, one should not forget that perfect (or Aristotelian) friendships are ideals. Even though it is hard to imagine that a human-robot friendship would achieve Aristotelian friendship, many of the human-human friendships in real life similarly do not realise that ideal either. However, this does not mean that imperfect friendships lack value in themselves, it simply means they are not as ideal as the perfect form (Danaher, 2019). Moreover, and on the bright side, there are advantages that imperfect forms of friendships have over perfect friendships. For example, there are greater risks entailed by Aristotelian friendship due to it being the ideal. When assuming a friendship is truly an Aristotelian friendship, one puts herself into a worse position when in actuality, the friend is manipulating and deceiving her than assuming that friendship actually takes the form of utility or pleasure friendship. Suppose one's gaming partner is lying about her education, career, and family, this would not affect the pleasure of gaming together or other social interactions, yet it would be completely different if their friendship was an Aristotelian one. Below, in the third chapter, I will sketch out the argument of Danaher (2019) on why it is possible that robots are virtue friends (i.e., realising mutuality) to humans when they are designed in specific ways. Accordingly, I will also introduce some risks and benefits of being in genuine human-robot relationships.

This chapter explained the phenomenon of anthropomorphism and its social and psychological reasons for the causation of such phenomenon in relation to robots. From a psychological point of view, the act of anthropomorphism is seen as a reason for feeling a greater sense of control in unfamiliar environments. Gaining control can be done by projecting a sense of humanness to non-human entities, to which the human subject becomes able to better recognise the anthropomorphised object with readily available storage of information that is the self. Relatedly, another psychological explanation is to imitate social connections, when the basic human need for social connection is lacking or absent, humans anthropomorphise robots and other technologies (because they are human-like and increasingly available in the modern technological world). The chapter also touched upon the Uncanny Valley Hypothesis (UVH) and the subsequent Dehumanisation Hypothesis. This is to demonstrate that anthropomorphising robots has many elements that reflect and relate to the human self, rather than robots' capabilities of engaging in intimate relationships with humans. Moreover, some historical background of social interactions between humans and machines are presented, such as the CASA paradigm, which denotes the first recognised human tendency to anthropomorphise and thereby act socially towards computers. Furthermore, the meaning of gendering robots implicates significant design insights in ethically creating robots that may benefit both humans and robots. Finally, the chapter closed with an introduction to human-robot friendship, starting with a normative section addressing the ethical concerns of befriending robots, and ending with the introduction to Aristotle's idea of the perfect friendship. In the following chapter, the perfect friendship of Aristotle will be applied to human-robot love and friendship, by incorporating the idea of Danaher (2019).

Chapter 3

How can intimate human-robot relationships be achieved

This chapter begins with a comparative analysis of the virtue model in human-human relationships and human-robot/technology relationships. Following that, the chapter adopts ideas from Danaher (2019) and argues that intimate relationships between humans and robots are, to a great extent, indeed possible, taking into account the concept of mutuality. However, for robots to engage in mutual intimate relationships with humans requires an investigation into whether AI can acquire phenomenal consciousness. This chapter then goes into presenting Searle's ideas of artificial consciousness, accompanied by some contemporary replies to Searle's ideas.

Aristotle's perfect friendship

In order to further investigate the possibilities of a human-robot friendship and assess whether it is truly possible, it is crucial to provide a description of what would a virtuous, Aristotelian friendship look like, and what are the necessary conditions to be satisfied in reaching such a perfect friendship. In the total of eight books of the *Nichomachean Ethics*, Aristotle (350 B.C.E) devoted two books to thoroughly discussing his understanding of the term 'Philia' (Tracy, 1979). Philia is most commonly translated as the 'highest form of love', and it is one of the four ancient Greek words for love. Within the context of *Nicomachean Ethics*, philia is typically translated to friendship, affection, and love (Liddell & Scott, 1940). Nonetheless, Tracy (1979) remarks that such standard translation is inadequate, and offers the following interpretation: "Aristotle sees philia, taken in the broadest sense of mutual attraction and attachment, as that which ties together, along with justice, every form of natural and conventional relationship among human beings" (p. 65). And this was also evident in the original text: "For in every association, we find mutual rights of some sort as well as philia" (Aristotle, 1159 b 26 f.). As such, it is important to bear in mind that although in the rest of the text I will refer to philia as friendship, the term also contains a broader sense of love, affection and most considerably, mutuality.

For Aristotle, friendship essentially encompasses virtues; friendship is a virtue or it implies virtue (Cocking, 2014; Danaher, 2019; Tracy, 1979). Aristotle's perfect friendship is distinctively valuable because of his claim that the perfect friendship exists between two (or more) who are (near) perfect in virtue, and this has been widely interpreted as "mutual recognition of one another's virtue that grounds and characterises the kind of love and respect found in ideal friendship" (Cocking, 2014, p. 85). Accordingly, appreciating each other's virtue becomes central to engaging in ways of intimate, deeply affectionate, trustful, and loyal friendships that are deemed ideal and perfect (Cocking, 2014). In striving to achieve a perfect friendship a number of conditions are needed to be satisfied for Aristotle, and Danaher (2019) selects four conditions that he finds most relevant and most central to a virtuous friendship (see table 1. below).

Authenticity/ honesty condition	*All participants involved in the friendship are honest with each other. Participants present themselves to each other as their true selves. Participants are not selective, duplicitous or manipulative.*
Equality condition	*All participants within the friendship perceive themselves as equals, that is, one party does not think they are superior to the other, since this could hinder mutuality.*
Diversity condition	*All participants interact with each other in a diverse and varied set of circumstances, as this enables a greater degree of mutuality*

	(e.g., one may obtain a utility friendship with her hairdresser, but this friendship is only exercised when she goes to the hairdresser, therefore it cannot be a perfect friendship according to Danaher and Aristotle)
Mutuality condition	*The friendship contains mutual sharing of interests and values, that is: "shared values, interests, admiration and well-wishing between the friends" (p. 6). "This is the most obvious condition since it is built into the definition of the friendship" (Danaher. 2017, p. 3).*

Table 1. The virtue model; the most salient conditions that should be fulfilled to enter an Aristotelian friendship, according to Danaher (2019).

Aristotelian human-robot friendship: can humans and robots engage in perfect companionship? The answer to this question, on its very surface, would be an absolute 'no' for most people, mainly due to the lack of mutuality in a human-robot relationship. If the virtue model is adopted as the understanding of friendship in the case, it would be impossible for robots to fulfil the four conditions (see table 3. below), and thus robots cannot be virtue friends with humans (Danaher, 2019; Elder, 2017; de Graaf, 2016).

Authenticity/ honesty condition	*Robots do not have an authentic self, they are artificially made by humans and therefore inauthentic in terms of their views,*

	values, principles etc… They are honest to their "friend" insofar as they are designed to be — this would be considered as an act of deception, since authenticity are only exhibited in "outward performances" (Danaher, 2019, p. 7), while lacking the internal mental states that cause outward performances. If robots are programmed to engage in meeting the conditions of the virtue model (i.e., behavioural and performative acts), it is equivalent to "hiring an actor to be your friend" (Danaher, 2019, p. 7; Elder, 2015; Nyholm & Frank, 2017).
Equality condition	*Humans are the creators of robots, therefore humans are the masters and are able to command and set tasks to robots, this makes robots always subservient to humans (Danaher, 2019). There would be inequality within any human-robot relationship until robots become as powerful as humans (Danaher, 2019).*
Diversity condition	*"it is difficult for robots to meet the diversity condition. For the time 8 being, robots will have narrow domains of competence. They will not be general intelligence, capable of interacting across a range of environments and sharing a rich panoply of experiences*

	with us. They cannot really share a life with us" (Danaher, 2019, p. 7).
Mutuality condition	*Robots do not have interests and values that stem from their robot selves. Robots' values are inputs through machine learning techniques by their programmers. Unlike humans, robots lack "inner mental life" (p. 7), which is crucial to mutuality (Danaher, 2019).*

Table 2. Popular philosophical arguments for rejecting the possibilities of robots as humans' virtue friends, according to Danaher (2017) and Danaher (2019).

Nonetheless, Danaher (2019) disagrees with the above arguments on the impossibility of virtue human-robot relationships according to Aristotelian ideals, and he provides two ways in which robots could satisfy the conditions of an Aristotelian friendship. Even though Danaher solely focuses on addressing human-robot friendship, these two approaches are equally applicable to the other intimate human-robot relationships that this thesis investigates. To understand the strengths of these approaches, one needs to first demarcate the kind of (im)possibility for robots to meet the four conditions. Danaher (2019) suggests that there is a significant difference between whether something is metaphysically impossible or technically impossible. Danaher (2019) claims that equality and diversity conditions are only technically impossible, meaning that it is metaphysically possible. On the other hand, authenticity and mutuality conditions are somewhat more robust to defend the view that robots would never achieve virtue friendship with humans because they are arguably both metaphysically and technically impossible (Danaher, 2019). To explain:

> *Equality is a function of one's powers and capacities and whether a robot is equal to a human with respect to its powers and capabilities is going to be dependent on its physical and computational resources, both of which are subject to technical innovations.* (Danaher, 2019, p. 8)

Likewise, the diversity condition is also currently impossible due to technical incompetence and constraints instead of metaphysical ones:

> *Whether a robot can interact with you across a diverse range of life experiences depends on its physical and computational dexterity (can it respond dynamically to different environments? can it move through them?) which is again subject to technical innovation.* (Danaher, 2019, p. 8)

On the contrary, authenticity and mutuality conditions are most likely to be dependent on consciousness and self-consciousness capacities in particular (Danaher, 2019). Granted that robots are authentically "computational objects" (p. 8) and their intelligence being artificial, it is difficult to counter argue the metaphysical impossibility to overcome the authenticity and mutuality conditions (Danaher, 2019). In addition, the technical possibilities are also, at the current stage, too technologically primitive to be worth considering (Danaher, 2019).

According to Danaher (2019), the first way in which human-robot friendship could be argued to approximate an Aristotelian ideal is grounded in the performative or behaviourist approach. Such an approach argues that insofar as robots perform in ways that the human users think they are meeting the four conditions, it is redundant if robots are not metaphysically or objectively capable of meeting those conditions (Danaher, 2019). Consequently, insofar as humans perceive robots as sharing their values and interests, imagine that robots have a certain inner mental life (e.g., through robot anthropomorphism), and decide that robots are relatively equal to humans, then robots are indeed capable of building ideal relationships with humans.

The second argument is a future argument, where one can argue that though the conditions above cannot be met by current AI robots, future robots might be able to meet

them (Danaher, 2019). Indeed, AI advancements have been developing at a breakneck speed, and it is likely to imagine that, through human input, training and self-learning, AI robots will one day be able to develop more intricate and sophisticated mental architectures (Danaher, 2019). Eventually, a so-called artificial consciousness may be realised, and robots may exhibit authenticity in themselves, in ways comparable to humans. Artificial consciousness would imply that robots would have some kind of autonomy and free will, in which they are able to choose to engage in relationships with humans where there is a mutual sharing of interests and values (I will discuss more artificial consciousness and its possibilities in the later section). All in all, when the aforementioned possibilities were to become actual capabilities in the future, robots would surely be able to interact with humans in diverse circumstances, in which interactions, and subsequently, relationships with humans would be more dynamic and less static. This argument may be deemed overly simplistic and generalisable, but mimicking nature has always been tried and done by engineers. Think of aeroplanes that mimic birds, or computers imitating human logical capabilities. All in all, as the great natural philosopher Plinius the Elder (70 CE) once ardently wrote: *Quam multa fieri non posse, priusquam sint facta, judicantur*! (*How many things, that were held to be impossible, have been accomplished!*).

Artificial consciousness: achievable or preposterous?

For robots to achieve mutuality in love and friendship with humans cognition is required, in particular, an aspect of cognitive functioning that is called consciousness.[19] The most valid way of realising authentic, meaningful human-robot companionship would be for robots to become conscious of themselves. Though intelligent robots have been readily available on the market in various forms in recent decades, conscious robots still seem like a far-fetched futuristic fantasy. If engineers succeed in replicating consciousness, which is

[19] There are two ways of understanding the relationship between consciousness and cognition. One sees consciousness as an aspect of cognitive functioning, or as cognitive in nature, the other sees it as distinct from any sort of cognitive functioning (Brown, 2014). I adopt the first understanding, which I find it more logical and more evidential because there are more concrete research and literature on the former (see e.g., Chalmers, 2017; Cohen & Dennett, 2011; Rosenthal, 2008). The latter is an assumption opposite of the first, and lacking evidence right now in literature.

arguably the most private and intimate feature of humans, then a fundamental question arises: what would it mean if there were artificial consciousness, or conscious robots? Not only would there be gigantic scientific progress, and would this have unprecedented social and technological implications and consequences, but also robots and humans would be significantly less different to each other. However, the more relevant question is: how realistic is it for artificial consciousness to be engineered?

In the past decades, an extensive amount of literature on the topic of artificial consciousness has emerged in the fields of neuroscience, philosophy of mind, as well as computer science and engineering. Before delving into the subject of artificial consciousness in robots, it is important to have a clearer understanding of consciousness. Searle (2013) provides a common-sense definition of consciousness as all those states of feelings, sensations, or awareness one goes through day in and day out. It begins in the morning when one wakes up from a dream, and it goes on all day until one falls asleep, dies, or becomes unconscious for other reasons. Scientifically, all conscious states are caused by lower-level neurobiological processes, and they are realised as higher-level or system features in the brain (Searle, 2013). According to Searle (2013), consciousness is a biological phenomenon that is inherent to most mammals, and perhaps other species. As such, the biological phenomenon that is present in known conscious beings like humans is absent in the design and engineering of robotics. This has been always a popular argument for the impossibility of computers being conscious, namely that consciousness stems from biological processes. Another argument that denounces the possibility of understanding the science of consciousness goes like the following: science is striving to be objective, and consciousness, on the other hand, is always subjective. Therefore, a science of consciousness simply cannot exist, and scientifically measuring and applying consciousness to computers is even more far-fetched. But is this philosophically correct?

Contrastingly, Dennett (1981) and Dennett (1991) give a positive account of consciousness by presenting the multiple draft model. The multiple draft model denies consciousness as

phenomenal or subjective and holds that one's thoughts and ideas are simply 'drafts' which are at various stages of editing. These drafts can be understood as different narrative fragments of the brain, and they do not necessarily go to a specific central processing unit. Instead, they may come together to serve particular purposes or functions (Dennett, 1991; Chua, 2017). For Dennett (1991), consciousness is purely functional, meaning that it consists of syntax only. In this case, consciousness can be imagined as a web of discourse, where the self is understood as the "centre of narrative gravity" (Chua, 2017, p. 48).

> *If the self is "just" the Centre of Narrative Gravity, and if all the phenomena of human consciousness are explicable as "just" the activities of a virtual machine realized in the astronomically adjustable connections of a human brain, then, in principle, a suitably "programmed" robot, with a silicon-based computer brain, would be conscious, would have a self. More aptly, there would be a conscious self whose body was the robot and whose brain was the computer. This implication of my theory strikes some people as obvious and unobjectionable. "Of course we're machines! We're just very, very complicated, evolved machines made of organic molecules instead of metal and silicon, and we are conscious, so there can be conscious machine.* (Dennett, 1991, as cited in Chua, 2017, p. 48)

This understanding of consciousness is a branch of computer functionalism, where the underlying idea is that the way a human thinks is similar to the way a computer programme functions. In this light, Dennett (1991) believes that conscious robots are eventually possible, unlike Searle (1981), and I will discuss this throughout this chapter. Dennett (1991) develops this argument by claiming that through more and more understanding of the ways functions and computations undergo in the human mind, these understandings are used as knowledge to be able to programme these functions and computations into machines.

The scientific and everyday explanation of consciousness. Searle (1980) emphasises that the problem with objectivity and subjectivity in science is that it is systematically ambiguous between an epistemic sense and an ontological sense. To explain this, we need to distinguish between epistemic objectivity and ontological objectivity categorically. Epistemically, the distinction between objectivity and subjectivity is between two types of knowledge claims. For example, the claim that Albert Einstein died in 1955 is epistemically objective, as it is a matter of objective fact. However, the claim that Albert Einstein is the greatest friend or the best bartender is a matter of opinion, hence epistemically subjective. On top of that, an overarching ontology brackets this epistemology; that is to say, there is a distinction regarding the mode of existence when it comes to subjectivity and objectivity. Things which are ontologically objective are those which exist regardless of whether anyone thinks otherwise. For instance, molecules, the process of osmosis, and the ocean all have modes of existence that are ontologically objective. On the other hand, feelings of pain, being in love, heartbreaks, and tickling only exist insofar as they are perceived and experienced by a subject. Accordingly, the latter list consists of things that are ontologically subjective. This distinction is crucial because, according to Searle (1980), many phenomena that are ontologically subjective belong to an account which is epistemically objective.

To return to the argument that there cannot be a science of consciousness (and hence no artificial consciousness), a fallacy of ambiguity now presents itself. Science is indeed epistemically objective since scientists strive for claims that can be established (as truth or not) independent of the attitudes of the creators and interpreters of the claim itself. However, the epistemic objectivity of the theory does not preclude an epistemically objective account of a domain that is ontologically subjective (Searle, 1980). Consequently, it is possible to have an epistemically objective science of consciousness, despite the phenomenon of consciousness is ontologically subjective.

Another important distinction that needs to be made is the difference between something that is observer-independent and something that is observer-relative.
Observer-independent things are things that are ontologically objective, where they exist regardless of whether people admit their existence or not. Here again, one can think of molecules, the process of osmosis, or the ocean. Contrastingly, observer-relative things only exist relative to observers, subjects, or users. They are things that have no meaning/ mode of existence in themselves; rather it is humans that give them meaning. Cash money is simply pieces of paper that humans store in their wallets; it is something that is observer-relative. The fact that makes cash money cash money is not the fact of its chemical composition. Rather, it is a fact about the attitudes that humans have towards it (Searle, 1980), these attitudes can be understood as the formation that is eventually realised in institutional practices. On the same note, love and friendship are also observer-relative; perhaps, the love between two lovers, no matter how spectacular and ideal their love may be, would only exist if it is perceived and understood by the lovers or other conscious agents.[20] Importantly, the distinction needs to be made because all observer-relative phenomena are created by the human consciousness, and therefore all observer-relative phenomena contain an element of ontological subjectivity (Searle, 1980). The scientific study of consciousness, hence, is a case where an epistemically objective science of a domain is rather observer-relative.[21] That is to say, despite science being observer-objective, a certain domain of science can still be observer-relative, and in this case, consciousness. Similar to consciousness in this context is the study of economics, which is also something ontologically subjective and observer-relative, but it is part of an epistemically objective scientific endeavour, according to Searle (1980).

[20] Some readers might be skeptical here and wonder: if love didn't exist in the first place, could there be an ideal version of love? To my understanding, an ideal of love can indeed exist in one's mind before the love comes into existence. It is usually what one hopes the love to be when it manifests.
[21] In an observer-relative sense, no matter how advanced a robot may be, it is not sufficient for thinking or being conscious. A highly advanced robot carries out computation that is exhibited as thinking, but computation is not a fact of nature, rather, it is a fact of the human interpretation (Searle, 1980).

Nonetheless, for Dennett (1991), consciousness does not include the key notion that Searle (1980) stresses as fundamental to the makeup of consciousness, namely qualia, and its ontological subjective properties. Whereas Searle (1980) views qualia as the unavoidable ontological subjective part of consciousness exclusive and unique to the self, it explains the impossibility of a conscious robot, Dennett (1991) denounces the existence of such ontologically subjective property. As such Dennett (1990) presents an entirely functional view of consciousness which exclusively consists of syntax, this differs from Searle (1980) as he emphasises the importance of both syntax and semantics. In other words, syntax alone does not resemble consciousness, it is the meaning (semantics) that allows consciousness to manifest.

Can AI understand? The famous thought experiment "The Chinese room" argues that despite the level of intelligence or the degree of human-likeness, AI cannot comprehend, understand, or have a mind of its own (Searle, 1980). The Chinese room argument goes roughly as follows: imagine person C is locked in a room. The room is full of Chinese characters, and contains a rule book for manipulating the characters. Unbeknownst to C, the room is a database, and the rule book is a programme. C is handed Chinese characters whilst staying in the room, unaware that these characters are questions written in Chinese. C looks into the rule book to check what she has to do, and she follows the instructions in the rule book by shuffling symbols and characters to give back sets of characters and symbols. The sets of characters she puts together are the correct answers to the questions in Chinese. Now suppose that over time C becomes extremely skilled at shuffling characters and symbols according to the given instructions from the rule book (i.e., C becomes skilled at writing the programme); her answers are indistinguishable from a native Chinese person. Even though she effortlessly passes the Turing test for understanding Chinese, she in fact does not understand a single character of Chinese. That being the case, it is impossible that C would come to actually understand Chinese, because C is locked in the Chinese room. The Chinese room is an analogy, whereby person C is metaphorically a computer system, and all the rules on which she operates are a computer programme. Most crucially, the programme itself is purely syntactical,[22] whereby it is unreservedly defined as a set of operations over syntactical elements (Searle, 1980). In a nutshell, if C (metaphorically the person in the room) does not understand the questions and the answers on the basis of implementing the programme, then on the same basis, neither does an AI system or digital computer. Computers are syntactical devices, and their operations are defined purely syntactically (Searle, 1980). This defines an essential difference between

[22] In philosophy of language, syntax is the study of the dynamics and make-ups of expressions that characterise them as components of varied linguistic categories, and 'well-formedness (Nearle, 1998). 'Well-formedness' can be understood as how these expressions belonging to these categories can be combined to construct greater units (Nearle, 1998). Syntactic properties have not only significantly contributed to the study of 'natural' languages, "a natural language or ordinary language is any language that has evolved naturally in humans through use and repetition without conscious planning or premeditation" (Langendoen, 1993, p. 1), such as English or Urdu), but also play an important role in the study of logic and computation (Nearle, 1998).

human intelligence and AI as the former requires not only syntax but also semantics[23], that is to say, it demands an understanding of the meanings of things.

The Chinese room thought experiment underlies two central principles that reject the possibility of robots being able to understand their human companions, and therefore disapprove of the possibility of mutuality in human-robot friendship and love. First: syntax is not semantics, and second: simulation is not duplication (Searle, 1980). The first principle demonstrates that understanding is distinct from following instructions that paint a facade of understanding. The second principle presents that in order to create human cognition in an AI system, not only do the behaviours of the human agent have to be simulated but the inherent cognitive processes that constitute those behaviours must also be duplicated (Searle, 2009). Following this logic, a robot companion is programmed to behave in all kinds of ways that resemble loving mutually in a relationship with a human, but the robot itself does not actually understand that it is showing love and being a good friend to the human. Moreover, even though all the behaviours of 'human' cognition are created and functioning on the robot, the robot still lacks the underlying cognitive process that accounts for its seemingly conscious behaviours.

While Searle (1980) proves the impossibility of a conscious machine, Dennett (1991) firmly stands with the idea that a conscious machine is possible. This is because Dennett and Searle hold different understandings of what a conscious robot entails. From Dennett's computation functionalism point of view of consciousness, the understanding of the entire functional system of humans' thinking process is sufficent to build a conscious robot. On the other hand, Searle contests this idea as he believes that functionalism is insufficient to create a conscious robot. Searle (1980) demonstrates his point through his Chinese room experiment and shows that although a computer programme may flawlessly mimic a language, environment or social order that resembles a conscious human, it lacks the

[23] In philosophy of language and cognitive science, semantics is the study of meaning in natural and artificial languages (Britannica, 2017).

understanding that a conscious human have in this case. Whereas Searle (1980) argues that consciousness is ontologically subjective and therefore is not independent of the human's biological construct, Dennett (1991) does not regard the biological construct as necessary to create a conscious robot, as he believes that consciousness can be created in complete materialism. In this section, I simply presented the accounts of Searle (1980) and Dennet (1991) without stating my stance on who is correct. This is because I understand these two theories as two distinct account of consciousness. Dennett (1991) holds the functionalist account of consciousness and Searle (1980) stresses the notion of understanding in consciousness by introducing the concept of qualia as the ontological subjective property that is unique to humans and other living things. To my knowledge, the two theories are largely incomparable, but they do present two different accounts of the future of artificial consciousness and conscious robots.

Information processing and consciousness

Information can be simply defined as the act of informing, whereby knowledge is transmitted by means of communication from a sender to a receiver that informs (the receiver) (Doyle, 2018). Information is an attribute that can be physically and mathematically explained for both humans and AI systems. In the common-sense understanding of information, it is stored in the brain of humans and in the storage of computers (Doyle, 2018). However, information simultaneously exists in any material, physical object, as well as in biological systems, such as genetic codes and cell structure, or the neurotransmitters messaging between and inside cells. "Contrarians" who denounce the world as informational often criticise this argument as limited and flawed, claiming that not everything is a message communicated with intended purposes. Nonetheless, in physics, scientists have established that structure is also information, as the structure in material objects is considered as things that can only be measured by an observer (Doyle, 2018). According to Doyle (2019), information philosophy understands material objects as:

> *... an information structure, from which the pure information can be abstracted as meaningful knowledge. In addition to structure, much of the information in living*

> *things consist of messages that are sent to invoke processes running in biological machines. Biological structures are often digital* (DNA, RNA, *proteins*) *and biological messaging greatly resembles a language, with arbitrary symbols, much like human language.* (p. 1)

Contrariwise, AI works by pattern recognition learning through interlacing large amounts of data set in the data that is set to be analysed, much of the information processes in humans (and other living things) "consists of messages that are sent to invoke processes running in the biological machines" (Doyle, 2019, p. 3). I argue that the ways of information processing (arguably, this also implies understanding) in humans and computers are highly comparable. Much like human languages, biological structures such as DNA, proteins, and RNA are often digital[24], and biological systems' communication adapts to arbitrary symbols (Doyle, 2019). Subsequently, to a great extent, these methods of messaging to various locations of the biological systems resemble the principal method of human communication, namely human language. Returning to the topic of this thesis, mutuality in a human-robot companionship can therefore be understood as the basis of an exchange of information between two types of intelligent systems, given that language is the method of communication that presents mutually shared values and interests.

Another objection to the Chinese room argument is that both humans and robots (AI programmes) are informational beings; in other words, they both function as beings that process information. Contrary to the common understanding of human cognition, this argument says that humans, like computers, also carry out computations. In fact, everything communicates with each other in the world we inhabit. Organisms (e.g., humans, animals, fungi and plants) can be understood as forms of intelligence which are posited in the respective information systems to communicate with each other via different means. In chapter two, it was argued that the uncanny feeling toward highly realistic humanoid robots stems from the grotesque uncertainty and the lack of understanding of what is a human. Correspondingly, humans have also not yet obtained a

[24] They are often digital as they are understood and explained through codes, according to Doyle (2019).

definition for life; that is, we do not know what life precisely is. Accordingly, the categorical distinction between death and life for beings and matters becomes somewhat artificial in the traditional understanding of the nature of being. If different life forms are to be understood as various kinds of information systems, then the idea that there are more forms of intelligence and information exchange would be beyond reasonable doubt (Hagedoorn, 2019). As such, programming languages that humans constructed are in a way understood by computers whereas human languages are what throughout history humans invented to communicate. When humans set commands to the computer through programming languages, I argue that this can be seen as a form of communication whereby information is exchanged between humans and robots. In other words, a mutual understanding is arguably at play here.

(Artificial) Intelligence and (artificial) consciousness

Intelligence is intricately related to consciousness (Gamez, 2020). Even though intelligence is a property that is purely functional, the property belonging to the conscious state is still uncertain for both humans and AI systems (Gamez, 2020). Searle, 1980 introduces two kinds of intelligence that are found in the world, namely observer-relative and observer-independent intelligence. The intelligence found in humans, dogs, or dolphins is intrinsic and observer-independent, and therefore they are comparable, and it is highly possible that dogs love[25] their owners in a similar way that the owners love their dogs (Searle, 1980). On the other hand, the computer is also highly intelligent, but the intelligence in computers is non-intrinsic and observer-relative. In the observer-relative sense, anything is a computer if one ascribes a computational interpretation (Searle, 1980). For instance, when one measures the speed of an object falling down, one is carrying out a computation. That is, one computes (calculates) using elementary mathematics. According to Searle, in the observer-relative sense, being a computer is not an intrinsic feature of an

[25] To my understanding, intelligence facilitates the ability to love, particularly in the dog and owner example. Because dogs have a higher intelligence than other animals (say fish or rodents), they are able to perceive and understand the affections they receive from their owners. Subsequently, they understand it is safe to be around their owners and feel comfortable and love their owners back.

object. Rather, it is a feature of the human interpretation of the physics of a phenomenon. This showcases that how much a robot is able to love its human back is highly personal, meaning that it depends on how much love the person perceives from the robot, supporting the performative claims of Danaher (2019).

Can AI (actually) think? If the definition of a machine is a physical system capable of performing certain tasks, then humanoid robots (i.e., AI) can indeed think, given that humans would also be understood as one kind of machine following this logic. Thinking is a biological process occurring in the brain in the form of certain complex but insufficiently understood neurobiological processes (Searle, 1980). In order to think, one needs to have a brain or something of equivalent causal powers to the brain. Take the simple equation 1 + 1 = 2, any conscious agent who is capable of carrying out this simple computation is both a computer and capable of thinking (Searle, 1980). This relates back to the performative and functional argument of Danaher (2019) stated above. Since humans and computers are both processing information and carrying out computations, a mutual understanding is arguably formed when humans act like they understand each other.

Nonetheless, unlike consciousness, computation is part of human interpretation, whereas consciousness exists outside of one's perception, and therefore is part of the world (not just the internal mental world of the human mind). All in all, Searle, 1980 argues that though scientific knowledge enables us to create an artificial mind, just like how an artificial heart can be produced, scientists are only able to simulate the mind yet incapable of duplicating it. As such, it is insufficient to have a model of consciousness to be able to run a consciousness by itself, it is only a model of consciousness. The computation is defined purely formally and syntactically, and therefore cannot be sufficient for thinking and any cognitive processes due to the lack of the element "understanding". [26]

[26] As the final chapter has come to an end, I would like to state that, so far, I presented these ideas of consciousness and AI in a purely descriptive manner. I think computation does not equal to thinking, like how John Searle advocated. Nonetheless, there are also weaknesses of the Chinese Room argument as I have presented earlier.

In the final chapter of this thesis, I sketched out love and friendship between humans and robots/technologies when applied to the Aristotelian ideals of friendship. For mutuality to be at play, the performative/ behaviourist approach would be needed to view the functions of companion robots. To have a robot and a human mutually valuing a friendship or relationship is not metaphysically impossible nor technically impossible, and thereby it is a matter of time before humans and robots can become genuine companions, even to the standard of a virtual friendship. Nonetheless, for robots to truly be a friend or lovers like another human requires robots to acquire consciousness and therefore "be" human. This topic encompassing strong AI and artificial consciousness is still much up to debate. Nevertheless, I find Searle's Chinese argument quite compelling in pointing out that despite a robot might act out very similar to how a human does, it does not mean that consciousness is duplicated. This can be summarised by Searle's famous line: that syntax is not semantics, and simulation is not duplication.

Embracing intimate human-robot relationships

This thesis has presented three main areas of discussion: 1) current philosophical understandings of what constitutes genuine love and friendship in human-human relationships. And more importantly, how, to love and be loved, contrasting popular beliefs, is partially guided by reasons, (partially irrational), and is essentially a moral phenomenon that calls away overpowering self-concern. 2) Anthropological and psychological explanations of why some humans display the tendency to experience being in (platonic or romantic) love towards robots, desiring or claiming that genuine love and friendship between humans and robots do exist despite popular controversies, justified by the theories of philosophy of love provided by 1). And 3), the performative/functional argument from the philosophy of technology on how mutuality can be accomplished in intimate

human-robot relationships, as well as the question in the philosophy of mind as to whether AI robots are capable of achieving consciousness to therefore part-take in authentic mutuality as it is arguably implied by any genuine, intimate relationship.

I argue that intimate human-robot companionships authentically exist and should be advocated and encouraged for the following reasons. First, as Jollimore (2011) showcases that, to love is to pave a path of seeing that involves bestowing charitable attention on a loved one, I argue that engaging in a loving relationship with robots helps to cultivate and enlarge human's ability to socialise, exhibit compassion, and love in more meaningful and ethical ways. This argument is also supported by Peeters and Haselager (2019) who criticise that most research on, for instance, sexbots, puts an unjust amount of focus on their use and interaction. They suggest that companion robots should be looked at in a non-instrumental way and that the study of human-robot companionship can yield a better understanding of ways humans think about love, friendship and sex. This is problematic as an overemphasis on interaction often leads to a neglect of attention on human-robot relations and other philosophical and scientific perspectives. Linking back to the uncanny valley argument presented by Aydin (2019), I argue that through the practice of genuine human-robot companionship, we would simultaneously learn more about ourselves, be it on an individual or collective level. This refers back to the UVH and Aydin (2020)'s philosophical analysis of the phenomenon.

Subsequently, I suggest that (companion or sex) robots could have substantial impacts on humans' moral character. In fact, if we move beyond the uni-directional, instrumental view, robots can teach us to live more fulfilling lives by enriching our social environment. For example, when meetings with patients, research and experiments are being systematically carried out in therapeutic settings, those who experience social difficulties can learn about ways to make friends and be more comfortable in social environments in practice sessions with trial companion bots. Peeters and Haselager (2019) suggest that virtue ethics may be a suitable framework to assess the implications of sex robots for human and robot relations,

as they r look into sex robots from a non-instrumental way, they believe that most research on sex robots gives too much unjust attention on the robot use and interaction, whereby if the matter is investigated more in-depth, one can argue that sex robots allow one to better understand the way humans think about love and sex. As such, robots can also potentially train some humans to interact with consent in romantic situations (Peeters & Haselager, 2019), cultivating socially desirable traits such as empathy and compassion to acknowledge it is not always about themselves, as love in a large sense means to give in a selfless way.

Good human-robot relationships: mutually beneficial. 'Healthy', intimate human-robot relationships do not exclusively benefit humans, the AI robots in these dynamics receive advantages too, respectively.[27] Relating back to the Lotbot case in chapter two, it is evident that the manner of talking to AI not only teaches AI to speak in the same manner, but humans are also greatly influenced by the manner in that AI talks back to them. If humans treat their robot companions in a compassionate, empathetic, and understanding way, AI registers and learns to behave in similar ways. In other words, humans and robots influence each other in behavioristic ways. Suppose humans take the initiative to behave kindly and lovingly toward AI robots. In that case, a virtuous cycle appears as AI would learn to behave by imitating humans and subsequently such a positive, social behaving manner is reinforced in humans. Again, this would result in humans treating each other more kindly too, in a way by practising kindness with robots. As such, it is possible to avoid AI robots 'gone wrong' cases such as Tay, the Microsoft AI who was corrupted by users on Twitter and became disturbingly racist and sexist in less than a day after she was launched (Vincent, 2016).

If humans ought to strive for the Aristotelian ideal of love and friendship, then it is only possible to take initiative on the human side. As I have attempted to argue throughout this thesis, intimate human-robot companionship is less encompassing the robots and is more

[27] This is meant in a metaphorical sense — we can shape AI into having 'better manners' for example, by treating them in a genuine way (regardless of whether they perceive it as genuine or not.

about the humans. It is ultimately about 'the human self', both as individuals and as a collective. AI robots are something humans invented for different reasons, and one of the reasons is, as this thesis has illustrated, to be caring human companions. And I argue that there is an abundance of potential social and individual merits to be gained when these relationships are studied more in-depth and less judged as freakish or wrong.

To my understanding, achieving friendship and love in human-robot partnerships goes hand in hand with ethically expanding into more areas of human-robot collaboration such as robot assistance in healthcare, security, economics, etc. Over the past few decades, these practices are continuing to progress but scandals exhibiting social discrimination also continue to make news headlines. A case in point is Amazon's recruiting robot learning to produce hiring decision outcomes biased against women (Ravi, 2021), or the U. S. border control AI security systems disproportionately targetting non-white ethnicities as suspects through false facial recognition matches (Marciano, 2019). Although this thesis acknowledges that AI failures like these are not entirely caused by robot misconduct (i.e., inter alia, design flaws and biased training samples and environment also contribute to the occurrence of AI discrimination), it suggests that maximising appropriate[28] human input during interactions/ training are crucial to minimising sexist and racist behaviours displayed in robots.

The aforementioned examples are evidence that there is an urgency to scrutinise ethical conduct regarding the human input of AI self-learning in the contemporary increasingly more AI-dependent world. As humans seek love and friendship with robots (viz., being more capable of engaging in more genuine and intimate relationships, and reducing the amount of loneliness), humans ought to take lead in demonstrating what constitutes good

[28] By appropriate I mean that there should be ethical guidelines designed for human-robot interactions and partnerships. For instance, humans who interact with AI robots should be made aware of the risks and consequences when feeding misdemeanours and malicious input. Since the possibilities of having robots as companions are growing, I suggest that modern humans should by default be educated to interact and live with robots in ethical ways. Of course, this may be a long-stretched proposition, given that this problem is still novel, but it is important to keep it in mind in the process of developing good human-robot companionships.

human-robot relationships, which is the groundwork for striving for perfect friendship or romance with robot companions. And in the effort to do so, I argue that some struggles and trial and error are morally permissible because it is only through downfalls that greater knowledge and outcomes can be obtained (this point is also mentioned in discussing Jollimore (2011)'s view on the epistemology of love). Abuse, exploitation, and neglect towards companion robots will inevitably happen, as they are historically illustrated in existing preliminary human-robot interaction trials like the Microsoft AI chatbot Tay and others. However, abuse is also everpresent in many human relationships, but for any relationship to be genuine and loving, all parties in the relationship need to understand that absence of abuse is a precondition.

Accordingly, I argue that achieving genuine human-robot companionship is a learning process, through which humans would not only learn to love robot companions but also learn to love their fellow humans, and more essentially, also to love themselves (as discussed in chapter one on Jollimore, 2011). To my understanding, when society accepts human-robot friendship and love as a norm, on the micro level, citizens would come to sense easier that it is in fact possible and appealing to live harmoniously with those who are mechanistically different (i.e., humanoid robots), that it is possible to love others who may seem uncanny and "not one of us". And on the macro level, citizens may learn to be more inclusive, more accepting and more tolerant towards other fellow humans who are different from them (viz., cultural background, upbringing, ethnicity, race, gender, political opinions, etc.). Historically, through social hardships and political endeavours to overcome traditions and conservative beliefs, different societies and cultures eventually became accepting and loving towards each other (to a certain extent). By the same token, to be able to accept human-robot companionship and to exercise love demands effort, especially in its infancy, the effort to ditch the prejudice that humans and robots cannot be friends or lovers because it is 'unnatural'.[29] In fact, "love isn't something natural. Rather it requires

[29] To argue that something is bad or untrue because it is 'unnatural' or something is good or true because it is 'natural' is known as the Appeal to nature argument, and is a logical fallacy (Elsher, 2022; Moore, 1922). This type

discipline, concentration, patience, faith, and the overcoming of narcissism. It isn't a feeling, it is a practice" (Fromm, 2013, p. 77).

Subsequently, by preventing robot abuse and establishing the importance to treat robots kindly, societal values are protected via humans' social engagement with social robots (Friedman et al., 2022; Darling, 2016). This is manifested in two ways, first, care and a lack of abuse are set as desirable examples to children on how to treat others (friends, partners, and also pets) (Friedman et al., 2022; Darling, 2016). And second, even as adults, the mental images of something alive and lifelike may be subconsciously blurred, and treating robots in ethical ways warrants maintaining the same attitudes and manners toward robot companions that sound adults carry towards their friends, partners or pets (Darling 2021). Heidegger was right that technology embodies a specific way of revealing the world, and humans hold power over reality in that revealing.[30] In essence, when humans truly love, their robot companions perceive love (and understand it in a humanly way or another) and reciprocate and expand love upon the collective of humans. For that reason, "love is a power which produces love" (Fromm, 2013, p. 30).

---

of argument is common but flawed because to say 'what is natural is good' is not a fact but an opinion, and it is often irrelevant. To describe something as 'natural' cannot alone come to the conclusion that it is superior, it needs factual evidence for support (Elsher, 2022). Moreover, 'natural' is a loaded term that is highly interpretative, everything can be considered 'natural' since it exists in space and time. To say something is unnatural implies that traditional values are always right, to a great extent. Computers and other modern technologies can be argued as unnatural way to operate the world. Homosexuality and polygamy can also be argued as unnatural, in how to love or who one loves.

[30] For Heidegger, reality is not absolute and the same with all time and culture. Rather, it exists in the relative sense (i.e., only in relations). Reality 'in itself' is incomprehensible for humans, humans are only able to understand the 'reality for humans'. This is because as soon as humans attempt to perceive or grasp, reality is no longer a thing 'in itself' (Heidegger, trans, 1977). Here again, I want to reinforce a final time that the actions of humans towards robots produce reactions relative to the human self, instead of the world or the robots who are receiving the actions. However, it is also through treating robots kindly that society receives kindness. Likewise, when humans treat robots in hateful ways, society also becomes hateful. Technology reveals and reflects human reality.

Conclusion

In closing, intimate human-robot relationships encompass the understanding of the human self. Anthropomorphism, specifically human attachment to humanoid robots, reveals distinct layers of human understanding. On a psychological level, anthropomorphised robots allow humans to control their social environment and fulfil their social needs in rather technologically alienated contemporary societies. Anthropologically, the UVH demonstrates the lack of knowledge of the human self in which an eerie emotional response spawns in the presence of highly realistic humanoid robots. Jollimore (2011) illustrates that love is something morally and epistemically relevant, love is neither wholly irrational nor fully rational, it is rather something in between. Moreover, Jollimore (2011) denotes the epistemic and moral value of love. If we apply this understanding to human-robot love and friendship, values can be realised and therefore intimate human-robot relationships should be encouraged for better self-understanding as well as learning to treat others in more moral ways. From a philosophy of technology perspective, even though mutually engaged human-robot love and friendship are thus far not fully technologically manifested, such mutuality is nonetheless not metaphysically nor technologically impossible as robotic research and development advance at the current pace.

Future research and limitations

In relation to future research, it may be fascinating to further explore artificial consciousness in relation to the design and development of companion robots. In the present thesis, artificial consciousness is investigated with respect to whether it can be achieved. However, the arguments predominately surround Searle's Chinese Room Argument, as such, counterarguments to Searle's ideas are only briefly discussed. Accordingly, this serves as the first limitation to this thesis, whereby even though having conscious robots as human companions is greatly helpful to the mutuality problem, the

thesis does not make an explicit link between (artificial) intelligence and consciousness and love and friendship.

On that note, the second limitation may be that intelligence, other human features such as emotion and the concept of love may be fundamentally distinct and thus incomparable (For more on this, see the Dick (2015) on the implementation of bonded rationality in AI). Nonetheless, I believe that to look into the possibilities of love and emotions of AI, the first step would indeed be investigating the intelligence embedded in AI.

The third limitation lies in the content of the thesis: not all research questions are answered. In particular, question 2a. 'Why do some humans desire intimacy with robots rather than other humans?' is not sufficiently answered in this thesis. It would be fruitful for future research to shed light on this question. This would need empirical research, because individual stances differ significantly on this topic and it would do injustice to generalise the psychological preferences of this arguably novel question.

A fourth limitation is that the relation between love and friendship is not explicitly distinguished and discussed. Although I attempted to defend that discussing romance and friendship in one thesis is not too problematic, I am also aware that perhaps the thesis would be more focused if I had chosen to only analyse one of the two topics.
A fifth limitation is that the section of 'embracing intimate human-robot relationships' is positioned arguably awkwardly. It could have been more coherently structured into the whole thesis. This section contains my own argument that I have build up from the theories and ideas examined in the three chapters. This section may be better fitted in a chapter of its own, instead of a small section right before the conclusion.

Finally, it is acknowledged that this thesis contains a generally positive narrative of AI and robot companionship, as well as a possible technological utopia tone. Such narration is arguably unique in the philosophical field but has simultaneously some weakness to it. As

such, the present thesis may be more argumentatively robust if more critical views of AI and robotic technologies are adopted, consequently, AI and robots can be understood through less tinted lenses. A dominant (despite maybe being uneducated) public opinion is that AI and other technologies would solve all current human problems (i.e., a manifestation of technology liberation), what many do not realise is that technologies such as AI actually is heavily dependent on aspects of materiality, this not only is produced through intensive labour but also generates unwanted consequences like ecological concerns. All in all, political, ethical, and financial consequences inevitably follow the current continuous technological advances accordingly, and this needs to be made aware at all levels of society.

Literature

Ahn, J., Kim, J., & Sung, Y. (2022). The effect of gender stereotypes on artificial intelligence recommendations. Journal of Business Research, 141, 50-59. doi:https://doi.org/10.1016/j.jbusres.2021.12.007

Amstrong, D. (September 17, 2013). Emotional attachment to robots could affect outcome on battlefield. University of Washington. Retrieved from https://www.washington.edu/news/2013/09/17/emotional-attachment-to-robots-could-affect-outcome-on-battlefield/

Anthony, S. (June 30, 2011). Lovotics, the new science of human-robot love. Extreme Tech. Retrieved from https://www.extremetech.com/extreme/88740-lovotics-the-new-science-of-human-robot-love
Aristotle, ., & Pakaluk, M. (1998). *Nicomachean ethics: Books VIII and IX*. Oxford: Clarendon Press.

Aydin. Extimate Technology- chapter 8. (2021). Technological uncanny as permanent structure of selfhood.

Aydin. Extimate Technology- chapter 8. Technological uncanny as permanent structure of selfhood.

Bendel, O. (2016). Sex Robots from the Perspective of Machine Ethics. *LSR*.

Bostrom, N. (2005). IN DEFENSE OF POSTHUMAN DIGNITY. Bioethics, 19(3), 202-214. doi:https://doi.org/10.1111/j.1467-8519.2005.00437.x

Brewer, K. B. (1996). Social identity theory. *Proceedings of the International Association for Business and Society*, 7, 1071-1082.

Britannica, T. Editors of Encyclopaedia (2017, May 1). semantics. Encyclopedia Britannica. https://www.britannica.com/science/semantics

Brown, R. (2014). Consciousness doesn't overflow cognition. Frontiers in psychology, 5. doi:10.3389/fpsyg.2014.01399

Bryson, J. J. and Kime, P. (1998). Just another artifact: Ethics and the empirical experience of AI. In *Fifteenth International Congress on Cybernetics*, pages 385–390.

Campbell, M., Hoane, A. J., & Hsu, F.-h. (2002). Deep Blue. Artificial Intelligence, 134(1), 57-83. doi:https://doi.org/10.1016/S0004-3702(01)00129-1

Camus, A. (1965). *The myth of Sisyphus, and other essays*. London: H. Hamilton.

Carter, C. S., & Porges, S. W. (2013). The biochemistry of love: an oxytocin hypothesis. *EMBO reports*, 14(1), 12–16. https://doi.org/10.1038/embor.2012.191

Carter, C. S., & Porges, S. W. (2013). The biochemistry of love: an oxytocin hypothesis. *EMBO reports*, 14(1), 12–16. https://doi.org/10.1038/embor.2012.191

Castello, O. (April 3, 2018). The Future of Love: Robot Sex and AI relationships. Retrieved from https://orge.medium.com/the-future-of-love-robot-sex-and-ai-relationships-3b7c7913bb07

Castello, O. (April 3, 2018). The Future of Love: Robot Sex and AI relationships. Retrieved from https://orge.medium.com/the-future-of-love-robot-sex-and-ai-relationships-3b7c7913bb07

Cerit, J., Dindarian, A., & Cilo-van Norel, N. (2020). Women, AI, and the Power of Supporting Communities: A Digital Gender-Support Partnership. Patterns, 1(1), 100009. doi:https://doi.org/10.1016/j.patter.2020.100009

Chalmers, D. (2017). The hard problem of consciousness. *The Blackwell companion to consciousness*, 32-42.

Cheetham M. (2017). Editorial: The Uncanny Valley Hypothesis and beyond. *Frontiers in psychology*, *8*, 1738. https://doi.org/10.3389/fpsyg.2017.01738

Cheok, A. D., Karunanayaka, K., & Zhang, E. Y. (2017). 192193Lovotics: Human–Robot Love and Sex Relationships. In P. Lin, K. Abney, & R. Jenkins (Eds.), *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (pp. 0): Oxford University Press.

Chua, J. M. (2017). Formulating Consciousness: A Comparative Analysis of Searle's and Dennett's Theory of Consciousness. *Talisik Undergraduate Journal of Philosophy*, 4(1), 43-58.

Chuah, S. H.-W., & Yu, J. (2021). The future of service: The power of emotion in human-robot interaction. Journal of Retailing and Consumer Services, 61, 102551. doi:https://doi.org/10.1016/j.jretconser.2021.102551

Chuah, S. H.-W., & Yu, J. (2021). The future of service: The power of emotion in human-robot interaction. Journal of Retailing and Consumer Services, 61, 102551. doi:https://doi.org/10.1016/j.jretconser.2021.102551

Cocking, D. (2014). Aristotle, Friendship and Virtue. *Revue internationale de philosophie*, 267, 83-90. https://doi.org/10.3917/rip.267.0083

Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in cognitive sciences*, 15(8), 358-364.

Danaher, J. (2019). The Philosophical Case for Robot Friendship. *Journal of Posthuman Studies*, 3(1), 5-24. doi:10.5325/jpoststud.3.1.0005

Darling, K. (2016). "Chapter 9: Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects". In *Robot Law*. Cheltenham, UK: Edward Elgar Publishing. Retrieved Jul 19, 2022, from https://www.elgaronline.com/view/edcoll/9781783476725/978178347672

Darling, K. (2021). The New Breed: What Our History with Animals Reveals about Our Future with Robots: Henry Holt and Company.

David, D., Hayotte, M., Thérouanne, P., d'Arripe-Longueville, F., & Milhabet, I. (2022). Development and validation of a social robot anthropomorphism scale (SRA) in a french sample. International Journal of Human-Computer Studies, 162, 102802. doi:https://doi.org/10.1016/j.ijhcs.2022.102802

Dege, M. (2014). Master-Slave Dialectic. In T. Teo (Ed.), Encyclopedia of Critical Psychology (pp. 1151-1156). New York, NY: Springer New York.

de Graaf, M. M. A. (2016). An Ethical Evaluation of Human–Robot Relationships. *International Journal of Social Robotics*, 8(4), 589-598. doi:10.1007/s12369-016-0368-5

Dennett, D. C. (1978). Why you can't make a computer that feels pain. In Brainstorms, pages 190–229. Bradford Books, Montgomery, Vermont. page numbers are from the 1986 Harvester Press Edition, Brighton, Sussex.
Dick, S. (2015). Of Models and Machines: Implementing Bounded Rationality. *Isis*, 106(3), 623-634. doi:10.1086/683527

Dobzhansky, T. (1973). Nothing in Biology Makes Sense except in the Light of Evolution. *The American Biology Teacher*, 35(3), 125–129. https://doi.org/10.2307/4444260

Dobzhansky, T. (1973). Nothing in Biology Makes Sense except in the Light of Evolution. *The American Biology Teacher*, 35(3), 125–129. https://doi.org/10.2307/4444260

Douthat, R. (June 5th, 2015). It is only a matter of time until polygamy is socially accepted. Las vegas Sun. Retrieved from https://lasvegassun.com/news/2015/jun/06/its-only-matter-time-until-polygamy-socially-accep/

Duffy, B. R. (2003). Anthropomorphism and the social robot. Robotics and Autonomous Systems, 42(3), 177-190. doi:https://doi.org/10.1016/S0921-8890(02)00374-3

Elder, A. M. (2017). *Friendship, Robots, and Social Media: False Friends and Second Selves*: Routledge.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological science*, *17*(4), 311-318.

Epley, N., Waytz, A., & Cacioppo, J. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. Psychological review, 114, 864-886. doi:10.1037/0033-295X.114.4.864

Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. (2008). When We Need A Human: Motivational Determinants of Anthropomorphism. Social Cognition - SOC COGNITION, 26, 143-155. doi:10.1521/soco.2008.26.2.143

Fedock, R., Kühler, M., & Rosenhagen, R. (Eds.). (2020). Love, Justice, and Autonomy: Philosophical Perspectives (1st ed.). Routledge. https://doi.org/10.4324/9780429323997

Fedock, R., Kühler, M., & Rosenhagen, R. (Eds.). (2020). Love, Justice, and Autonomy: Philosophical Perspectives (1st ed.). Routledge. https://doi.org/10.4324/9780429323997

Fong, T., Dautenhahn, K., Scheutz, M., and Demiris, Y., editors (2008). Governing lethal behavior: Embedding ethics in a hybrid deliberative/reactive robot architecture, New York. ACM.
Fosch-Villaronga, E., & Poulsen, A. (2020). Sex care robots: Exploring the potential use of sexual robot technologies for disabled and elder care. Paladyn, Journal of Behavioral Robotics, 11(1), 1-18. doi:doi:10.1515/pjbr-2020-0001

Fosch-Villaronga, E., & Poulsen, A. (2020). Sex care robots: Exploring the potential use of sexual robot technologies for disabled and elder care. Paladyn, Journal of Behavioral Robotics, 11(1), 1-18. doi:doi:10.1515/pjbr-2020-0001

Fosch-Villaronga, E., & Poulsen, A. (2020). Sex care robots: Exploring the potential use of sexual robot technologies for disabled and elder care. Paladyn, Journal of Behavioral Robotics, 11(1), 1-18. doi:doi:10.1515/pjbr-2020-0001

Frankfurt, H. G. (2004). The Reasons of Love: Princeton University Press.

Fromm, E. (1989). *The art of loving*. New York: Perennial Library.

Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a stronger CASA: Extending the computers Are social actors paradigm. Human-Machine Communication, 1, 71-86. https://doi.org/10.30658/hmc.1.5

Gamez, D. (2020). The Relationships Between Intelligence and Consciousness in Natural and Artificial Systems. Journal of Artificial Intelligence and Consciousness, 07(01), 51-62. doi:10.1142/s2705078520300017

Garland, A. (Director). (2014). Ex Machina. [Film]. Film for DNA films.

Gehlen, A. (2003). A philosophical–anthropological perspective on technology. In R. C. Scharff & V. Dusek (Eds.), Philosophy of Technology: The Technological Condition: An Anthology (pp. 213--220): Blackwell.
Geysegom (1997). Samenlevingscontracten 1997: focus op het buurtniveauPreteskt. (12): 17-19.

González-González, C. S., Gil-Iranzo, R. M., & Paderewski-Rodríguez, P. (2020). Human-Robot Interaction and Sexbots: A Systematic Literature Review. *Sensors (Basel, Switzerland)*, *21*(1), 216. https://doi.org/10.3390/s21010216

Guthrie, S. (1993). Faces in the Clouds: A New Theory of Religion: Oup Usa.

H, Bennett, "Love". (2021). *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.). https://plato.stanford.edu/archives/fall2021/entries/love/.

H, Bennett, "Love". (2021). *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.). https://plato.stanford.edu/archives/fall2021/entries/love/.

Hagedoorn, P. (May 15, 2019). Everything is Information. The fluid society. Retrieved from https://hagedoorn.org/en/everything-is-information/

Haladjian, H. H., & Montemayor, C. (2016). Artificial consciousness and the consciousness-attention dissociation. *Consciousness and Cognition*, 45, 210-225. doi:https://doi.org/10.1016/j.concog.2016.08.011

Han, Y. (2021). Do We Love For Reasons? Philosophy and Phenomenological Research, 102(1), 106-126. doi:https://doi.org/10.1111/phpr.12638

Heidegger, M., & Lovitt, W. (1977). *The question concerning technology, and other essays*.

Infotaris. (April 14, 2019). Samenlevingscontract in het Engels: free example in English. Retrieved from https://www.infotaris.nl/blogs/samenlevingscontract-engels/

Iturriaga, N., & Saguy, A. C. (2017). "I Would Never Want to be an Only Wife": The Role of Discursive Networks and Post-Feminist Discourse in Reframing Polygamy. *Social Problems*, 64(3), 333–350. https://www.jstor.org/stable/26505536

Jenkins, C. S. I. (2017). *What Love Is: And What It Could Be:* Basic Books.

Joe, W. Y., & Song, S. Y. (2019). Applying Human-Robot Interaction Technology in Retail Industries. International Journal of Mechanical Engineering and Robotics Research, Vol. 8, No. 6, pp. 839-844. DOI: 10.18178/ijmerr.8.6.839-844

Jollimore, T. (2011). *Love's Vision*. Princeton: Princeton University Press. https://doi.org/10.1515/9781400838677

Klein, J. (March 25th, 2021). Ethical non-monogamy: the rise of multi-partner relationships. BBC. Retreived from https://www.bbc.com/worklife/article/20210326-ethical-non-monogamy-the-rise-of-multi-partner-relationships

Krauss, N. (2005). *The history of love.* New York: W. W. Norton.

Langendoen, D. T. (1993). [Review of *Natural Language and Universal Grammar*, by J. Lyons]. *Language*, 69(4), 825–828. https://doi.org/10.2307/416893

Lee, J.-E., & Nass, C. (2010). Trust in Computers: The Computers-Are-Social-Actors (CASA) Paradigm and Trustworthiness Perception in Human-Computer Communication. Trust and Technology in a Ubiquitous Modern Environment: Theoretical and Methodological Perspectives, 1-15. doi:10.4018/978-1-61520-901-9.ch001

Lee, K. M., Park, N., & Song, H. (2006). Can a Robot Be Perceived as a Developing Creature?: Effects of a Robot Long-Term Cognitive Developments on Its Social Presence and People Social Responses Toward It. Human Communication Research, 31(4), 538-563. doi:10.1111/j.1468-2958.2005.tb00882.

Leyens, J.-P., Cortes, B., Demoulin, S., Dovidio, J. F., Fiske, S. T., Gaunt, R., . . . Vaes, J. (2003). Emotional prejudice, essentialism, and nationalism The 2002 Tajfel lecture. European Journal of Social Psychology, 33(6), 703-717. doi:https://doi.org/10.1002/ejsp.170

Liddell, H. G., Scott, R., Jones, H. S., & McKenzie, R. (1940). A *Greek-English lexicon.*

Maeng, A., & Aggarwal, P.K. (2018). Facing Dominance: Anthropomorphism and the Effect of Product Face Ratio on Consumer Preference. *Journal of Consumer Research*, 44, 1104-1122.

Mamdani, Z. (February 28, 2008). Idaho Evangelical Christian polygamists use Internet to Meet potential spouses. Deseret News. Retrieved from https://www.deseret.com/2008/2/28/20073463/idaho-evangelical-christian-polygamists-use-internet-to-meet-potential-spouses

Marciano, A. (2019). Reframing biometric surveillance: from a means of inspection to a form of control. *Ethics and Information Technology*, *21*(2), 127–136. https://doi.org/10.1007/s10676-018-9493-1

Mckeever, N. (2012). Love's Vision – By Troy Jollimore. Journal of Applied Philosophy, 29(1), 88-90.

Merlhiot, G., Mondillon, L., Méot, A., Dutheil, F., & Mermillod, M. (2021). Facial width-to-height ratio underlies perceived dominance on facial emotional expressions. *Personality and Individual Differences*, *172*, 110583. doi:https://doi.org/10.1016/j.paid.2020.110583

Moore G. E. (1922). *Principia ethica.* Cambridge University Press.

Mori, M., & MacDorman, K.F. (2017). The Uncanny Valley: The Original Essay by Masahiro Mori-IEEE Spectrum.

Mori, M., & MacDorman, K.F. (2017). The Uncanny Valley: The Original Essay by Masahiro Mori-IEEE Spectrum.

Mugar, G. (March 27, 2019). Why It's Important to Say Please and Thank You to Robots. Retrieved from https://www.ideo.com/blog/why-its-important-to-say-please-and-thank-you-to-robots

Neal, M. (September 18, 2013). Are Soldiers Getting Too Emotionally Attached to War Robots? Vice. Retrieved from https://www.vice.com/en/article/jppy8g/are-soldiers-getting-too-emotionally-attached-to-war-robots

Neal, M. (September 18, 2013). Are Soldiers Getting Too Emotionally Attached to War Robots? Vice. Retrieved from https://www.vice.com/en/article/jppy8g/are-soldiers-getting-too-emotionally-attached-to-war-robots

Neale, S.(1998). Syntax. In The Routledge Encyclopedia of Philosophy. Taylor and Francis. Retrieved 28 Jun. 2022, from https://www.rep.routledge.com/articles/thematic/syntax/v-1. doi:10.4324/9780415249126-W044-1

Nelson, J. (2014). Positivism. In T. Teo (Ed.), *Encyclopedia of Critical Psychology* (pp. 1437-1444). New York, NY: Springer New York.

Newport, F. (June 26th, 2020). Understanding the increase in moral acceptability of polygamy. Gallup. Retrieved from https://news.gallup.com/opinion/polling-matters/313112/understanding-increase-moral-acceptability-polygamy.aspx#:~:text=In%20short%2C%20there%20has%20been,20%25%20morally%20acceptable%20in%202020.

Nyholm, S., & Frank, L. (2017). From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible? In Danaher and McArthur, Robot Sex: Social and Ethical Implications, Cambridge, MA: The MIT Press.

Nyholm, S., & Frank, L. (2017). From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible? In Danaher and McArthur, Robot Sex: Social and Ethical Implications, Cambridge, MA: The MIT Press.

Nyholm, S. (2020). Humans and Robots: Ethics, Agency, and Anthropomorphism.

Nyholm, S. (2020). Humans and Robots: Ethics, Agency, and Anthropomorphism.

O'Leary, A. (1 December, 2020). Mirror. Bodybuilder marries sex doll after whirlwind romance during coronavirus lockdown. Retrieved from https://www.mirror.co.uk/news/world-news/bodybuilder-marries-sex-doll-after-23097953

Phillips, J., Misenheimer, L., & Knobe, J. (2011). The Ordinary Concept of Happiness (and Others Like It). *Emotion Review*, 3(3), 320–322. https://doi.org/10.1177/1754073911402385

Plato. (1989) The Symposium. Indianapolis, IN: Hackett Publishing Company.

Prescott, T. J., & Robillard, J. M. (2021). Are friends electric? The benefits and risks of human-robot relationships. iScience, 24(1). doi:10.1016/j.isci.2020.101993

Reuter, K. (2019). Dual character concepts. Philosophy Compass, 14(1), e12557. doi:https://doi.org/10.1111/phc3.12557

Reuter, K. (2019). Dual character concepts. Philosophy Compass, 14(1), e12557. doi:https://doi.org/10.1111/phc3.12557

Rips, L. J. (1975). Inductive judgments about natural categories. Journal of Verbal Learning & Verbal Behavior, 14(6), 665-681. doi:10.1016/S0022-5371(75)80055-7

Roomba. Robots your guide to the world of robotics. (2022). Retrieved from https://robots.ieee.org/robots/roomba/

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia*, *46*(3), 829-840.

Samani, H., Cheok, A., Tharakan, M., Koh, J., & Fernando, O. (2010). A *Design Process for Lovotics* (Vol. 59).

Saunderson, S. (2021). Robots can be companions, caregivers, collaborators— and social influencers. Retrieved from https://theconversation.com/robots-can-be-companions-caregivers-collaborators-and-social-influencers-172215

Scaringi Law. (2022). Polygamy, Religion, and the Law. Retrieved from

https://www.scaringilaw.com/blog/2020/january/polygamy-religion-and-the-law/#:~:text=For%20example%2C%20in%20some%20Islamic,multiple%20husbands%20and%20multiple%20wives

Schreier, J. (Director). (2012). Frank & Robot. [Film]. Hellowell House.

Scrubb, V. (August 11, 2016). Sexbots and Terminators: Exploring Gender in AI. Artificial Shane. Retrieved from https://shanesaunderson.com/2016/08/11/sexbots-and-terminators-exploring-gender-in-ai/

Searle, J. (2009). Chinese room argument. Scholarpedia, 4(8):3100.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, 3(3), 417-424.

Semin, G. R., & Fiedler, K. (1988). The Cognitive Functions of Linguistic Categories in Describing Persons: Social Cognition and Language. Journal of Personality and Social Psychology, 54, 558-568.

Shakespeare, W. (1993). *Romeo and Juliet*. Dover Publications.

Simon, M. (October 3, 2018). It's Time to Talk About Robot Gender Stereotypes. Wired. Retrieved from https://www.wired.com/story/robot-gender-stereotypes/

Sparrow, R., & Sparrow, L. (2006). In the hands of machines? The future of aged care. Minds and Machines, 16(2), 141-161. doi:10.1007/s11023-006-9030-6

Spatola, N., & Wudarczyk, O. A. (2021). Ascribing emotions to robots: Explicit and implicit attribution of emotions and perceived robot anthropomorphism. Computers in Human Behavior, 124, 106934. doi:https://doi.org/10.1016/j.chb.2021.106934

Stinson, D., Cameron, J., & Hoplock, L. (2021). The Friends-to-Lovers Pathway to Romance: Prevalent, Preferred, and Overlooked by Science. Social Psychological and Personality Science, 13, 194855062110269. doi:10.1177/19485506211026992

Szczuka, J.M., & Krämer, N.C. (2016). Influences on the Intention to Buy a Sex Robot - An Empirical Study on Influences of Personality Traits and Personal Characteristics on the Intention to Buy a Sex Robot. *LSR*.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup

behaviour. *European Journal of Social Psychology*, *1*(2), 149-178. doi:https://doi.org/10.1002/ejsp.2420010202

Tracy, T. (1979). Perfect Friendship in Aristotle's "Nicomachean Ethics." *Illinois Classical Studies*, 4, 65–75. http://www.jstor.org/stable/23061134

Trovato, G., Lopez, A., Paredes Venero, R., & Cuellar, F. (2017). Security and guidance: Two roles for a humanoid robot in an interaction experiment.

Urquiza-Haas, E. G., & Kotrschal, K. (2015). The mind behind anthropomorphic thinking: attribution of mental states to other species. Animal Behaviour, 109, 167-176. doi:https://doi.org/10.1016/j.anbehav.2015.08.011

Veda. (February 15, 2022). Romantic Robots: Can Humans & Robots Love Each Other? Analytics Insight. Retrieved from https://www.analyticsinsight.net/romantic-robots-can-humans-robots-love-each-other/

Vincent, J. (March 24, 2016). Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. The Verge. Retrieved from https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

Vogels, E., Anderson, M. (May 8th, 2020). Dating and Relationships in the digital age. Pew Research Centre. Retrieved from https://www.pewresearch.org/internet/2020/05/08/dating-and-relationships-in-the-digital-age/

Wan, E. W., & Chen, R. P. (2021). Anthropomorphism and object attachment. Current Opinion in Psychology, 39, 88-93. doi:https://doi.org/10.1016/j.copsyc.2020.08.009

Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The Uncanny Valley: Existence and Explanations. *Review of General Psychology*, *19*(4), 393–407. https://doi.org/10.1037/gpr0000056

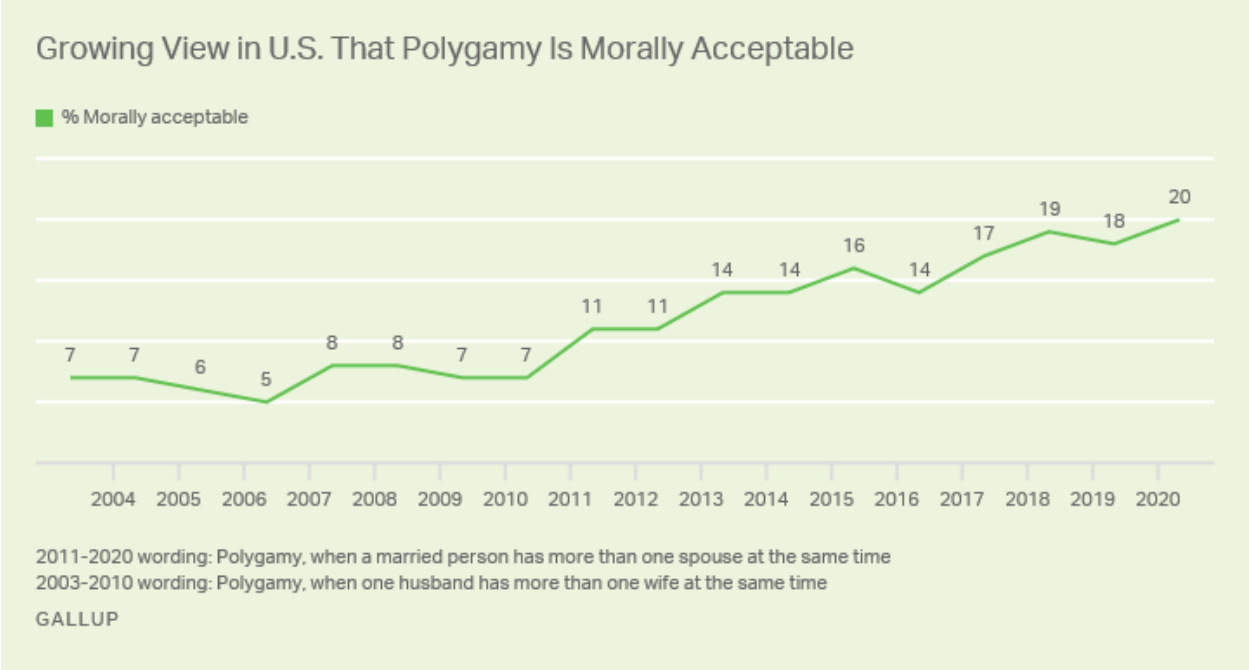
Weil, S. (1952). *Gravity and Grace* (Vol. 28): Routledge.

Wired. (June 16, 2003). The New Pet Craze: Robovacs. Retrieved from https://www.wired.com/2003/06/the-new-pet-craze-robovacs/

Appendix A

Graph of Polygamy's Moral Acceptability in the United States from 2004 to 2020

The line graph of Newport (2020) displays a fourfold increase in the public's acceptance of polygamy in the United States in two decades.



Appendix B

Image of facial width-to-height ratio and perceived dominance on facial emotional expressions (Merlhiot et al. 2021)

According to Wan and Chen (2021), evolutionary research demonstrates that impressions of dominance and aggression are perceived more strongly in human faces with "high face to width-to-height ratio (fWHR)" (p. 91).