philosophies

MDPI

Article

Happiness in the AI Age: Ricoeur and the Question of the AI Humanoid as the Technological Other

Anné Hendrik Verhoef [1] and Edmund Terem Ugar [2,*]

[1] School of Philosophy, Northwest University South Africa, Potchefstroom 2531, South Africa; anne.verhoef@nwu.ac.za
[2] Department of Philosophy, University of Johannesburg, Johannesburg 2092, South Africa
* Correspondence: edmundu@uj.ac.za

Abstract

In this paper, we examine the evolving conception of the "other" in relation to human happiness, drawing on Paul Ricoeur's philosophical account and empirical findings from the Harvard Study of Adult Development. Ricoeur situates happiness in three interrelated threads: individual fulfilment, friendship with those near to us, and just relations with distant others. The Harvard Study corroborates the significance of relationality for well-being, showing that strong social ties enhance longevity and life satisfaction. However, contemporary digitalisation and the proliferation of AI humanoid social robots challenge traditional notions of the "other." Individuals increasingly form "meaningful" attachments, emotional bonds, and even romantic relationships with technological artefacts, raising the question of whether these non-human entities can contribute to happiness in a Ricoeurian sense. While the above dynamics are now proliferating, we argue that AI and social robots cannot be considered as the "other" in the Ricoeurian sense. Although these technologies can be considered as a virtual other, we do not defend that position in the current paper. In this paper, we explore the tensions regarding the authenticity, moral status, and ethical implications of AI and social robots in relation to human happiness. We conclude by proposing a re-evaluation of relationality, moral consideration, and the ethical frameworks underpinning human–technology interactions in the pursuit of human flourishing and happiness in the Ricoeurian sense.

Keywords: happiness; paul ricoeur; social robots; relationships; other

1. Introduction

Recently, one of the world's longest longitudinal studies of adult life, the Harvard Study of Adult Development, found that strong community ties help people live longer and happier lives. In *The Good Life* [1], Robert Waldinger and Marc Schulz draw on nearly 80 years of data from 268 Harvard sophomores in the Harvard Study, which has been tracked since 1938, to bolster their argument. The study shows that close relationships best predict longevity, health, and life satisfaction. Such ties protect against discontent, delay physical and cognitive decline, and are more reliable predictors of well-being than wealth or status. A wider understanding of well-being as equated with happiness is argued for within the happiness sciences and analytic philosophy [2–4]. Happiness, however, is an area of broader philosophical and theological concern, and this current paper seeks to advance its understanding through Ricoeur's work.

Check for updates

Received: 6 March 2026
Revised: 21 April 2026
Accepted: 29 April 2026
Published: 25 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

The philosophical arguments on happiness presented by Paul Ricoeur, particularly in *Fallible Man* [5] and *Le Bonheur Hors Lieu* [6], resonate with the findings of the Harvard Study. In *Le Bonheur Hors Lieu*, Ricoeur advances three interrelated claims along the Aristotelian line of *eudaimonia* as a good, ethical life. First, happiness reflects the individual's aspiration towards a fulfilled life. Secondly, it is realised in friendship and proximity to others. Thirdly, it is expressed through justice and life within just institutions.

Ricoeur thus links the private pursuit of happiness to the role others play in enabling or frustrating it. The "other" includes those with faces—friends and loved ones—and the faceless, those distant strangers who nevertheless call for justice[1]. Happiness, then, may be located within the self, in intimate relationships, or in just relations with the wider community. The Harvard findings suggest that happiness is indeed found in relation to others, especially those near us (friends and family), and in our community. Yet the rise of digitalisation and AI complicate who and what may count as the "other", with a face or faceless, in the promotion of our individual happiness.

This paper, therefore, asks: (i) Is Ricoeur's conception of the "other" limited in the AI era? (ii) Can it be extended to non-human entities such as technological artefacts? What does this mean for happiness? These questions are increasingly urgent in light of evolving human–robot relationships, in which artefacts may supplement or even replace human companionship. For example, in their study, theorists such as Emont [7] presented data showing that elderly people in care facilities prefer care robots, while Suzuki et al. [8] showed that individuals with social anxiety preferred to interact with robots that lacked human resemblance. We engage some of these examples later in this paper. However, it is important to note that recent scholarship on human-technology relations highlights the social disruptions technology causes to our existing values and norms in relationships. For example, some scholarship explores romantic and sexual attraction to fictional or digital characters, sometimes described as fictoromantic or fictosexual orientations [9].

In Japan, there is a rising social recluse amongst over one and a half million individuals, often associated with the phenomenon of *hikikomori*, which has coincided with growing attachment to virtual companions [9]. Some individuals report that human relationships are stressful and instead turn to social robots, which are perceived as incapable of betrayal. A prominent example is Akihiko Kondo, who publicly married Hatsune Miku, a holographic device developed by Gatebox [9]. Thousands of individuals have reportedly sought similar symbolic unions. There are currently an estimated 3708 individuals who have applied to get marriage certificates through Gatebox with Hatsune Miku [10].

Such developments sharpen the question of what it means to be the "other". In these cases, the "other" is neither the intimate friend nor the distant stranger envisioned by Ricoeur, but a technological artefact. This paper interrogates whether such artefacts can meaningfully contribute to happiness in a Ricoeurian sense or be conceived as the "other" in the above sense. On the one hand, we present a sceptical hypothetical view of Ricoeur. We claim that these technologies may be able to contribute to human happiness and promote human flourishing, even though Ricoeur may gravitate towards the idea of the "other" as a human person rather than a social robot or AI. On the other hand, in light of the first argument, we make it clear (in the final section of this paper) that our position is that AI and social robots do not meet the criteria for becoming the "other". The point here is that AI and social robots may rightly contribute to human flourishing and happiness, but not from the position of the other in a Ricoeurian sense. Perhaps AI and social robots may be considered virtual others or quasi-others.

Before we proceed to the outline of the paper, we delimit our defended position. Although AI and social robots may qualify as partial, virtual, or quasi others, we do not advance this claim herein. We acknowledge that these technologies can foster human

flourishing and happiness, but not from the standpoint of the Ricoeurian other (human otherness). As we will show, they fail to satisfy the criteria for human otherness. These criteria are as follows: (i) we show that the technologies are simulations; (ii) their appeal to agency is but market hypes; (iii) they fail to satisfy the criteria for moral patiency; (iv) they do not have subjective experiences.

The paper proceeds as follows. First, we clarify the conception of happiness under consideration, positioning Ricoeur's philosophical account within the broader philosophical and religious interpretations. Secondly, we outline Ricoeur's framework and its relevance to the present paper. Thirdly, we examine tensions within this framework in light of contemporary human–technology relationships. Finally, we make our stance clear by presenting four arguments against the "otherness" of AI and social robots. We also indicate avenues for future research.

2. On Happiness: Our Delimitation

In philosophy, happiness is often understood as a mental state that produces joy or pleasure (cf., for instance, Haidt [3,11,12]). In some accounts, happiness is equated with well-being or flourishing (cf. [2,13]); in others, it is conceived as an ethical way of life [14], the absence of pain, or even good fortune. Across these approaches, happiness is typically presented as something we ought to pursue in our daily lives. Haidt [11], for example, argues that genuine happiness requires intentional activity. In such views, happiness is treated as immanent rather than grounded in a supernatural source.

By contrast, transcendental accounts of happiness are rooted in religious traditions and closely tied to the notion of God. Here, happiness is derived from communion with God. As Venter [15] suggests, promoting happiness requires Christian theology to "retrieve an acute sense of the Triune God, and to explore the potential of this symbol to contribute to human flourishing in a specific and concrete context."

From a Christian perspective, happiness is associated with transcendence. A distinction is often drawn between earthly and heavenly (eschatological) happiness. Heavenly happiness refers to eternal life with God after death. In this context, St Augustine of Hippo maintains that happiness arises from the joy experienced when one's deepest desires are fulfilled [16]. In *Happy Life*, Augustine likens the human search for happiness to sailors seeking a harbour: after roaming and encountering misfortune, they long for a place of rest [16]. This implies that human beings, like sailors, can only find happiness when they anchor their lives with God.

Augustine argues that happiness must not be sought in fleeting or contingent goods, but in what is permanent [16]. For him, that permanence is God. A happy life requires understanding God and cultivating the soul. Happiness is characterised by wisdom and truth, both of which stem from God. A happy soul participates in divine truth and delights in God [16].

Both immanent and transcendental accounts of happiness face criticism. The above theorists contend that philosophical or immanent approaches often reduce happiness to individualistic, hedonistic mental states, as seen in some strands of positive psychology, where concepts such as meaning, belonging, or hope lack clear definitions. Conversely, theology is criticised for advancing an exclusive conception of happiness accessible only to believers and for overemphasising the afterlife. As Verhoef [17] argues, theology maintains that "the highest possible form of happiness is thus only available for a religious person," while philosophical accounts are deemed inferior for relying solely on immanence.

Nevertheless, Verhoef [17] identifies an overlap between immanence and transcendence in Paul Ricoeur's philosophical anthropology (*Fallible Man*). Ricoeur resists reducing

happiness and human beings to mere causal entities devoid of complexity and mystery. Ricoeur's significance lies in the bridge he constructs between happiness and human nature.

It is within this Ricœurian framework that we situate our own account of happiness. Our aim is to clarify the blurred line between what it means to be the "other" and how this relates to happiness. Here, we limit the notion of the other to the human other and the technological artefact. The central question is whether the latter can meaningfully be regarded as an "other" in light of contemporary developments in human–technology interaction and human happiness. To this end, we explicate Ricoeur's understanding of happiness and the role of others in its attainment.

3. Paul Ricoeur on Happiness and the Role of the 'Other'

As noted in the introduction, Ricoeur's principal philosophical reflection on happiness appears in *Le Bonheur Hors Lieu* [6]. In this work, Ricoeur identifies three interrelated threads in the search for happiness. First, happiness concerns the individual's desire for a fulfilled life. Secondly, it is found in friendship and in those who are close to us. Thirdly, it is realised in justice and in living with others within just institutions. Ricoeur says happiness is to be found within these three threads, but not completely in only one of them, or in all of them together. Happiness remains 'hors lieu', something elusive and fragile, but that doesn't mean we should give up on finding happiness. We will return to this point at the end of the article.

We now offer a detailed analysis of the three happiness threads Ricoeur identified. Drawing on Ricoeur's text, we clarify the role of the "other" in the constitution of happiness. However, before getting into the notion of the "other" as it appears in Ricoeur's aforementioned work, a popular understanding of happiness in Ricoeur's hermeneutic and dialectic method is that happiness is comprehensive and holistic. In his work, Ricoeur offers a holistic understanding of happiness. Ricoeur constructs happiness as something linked to chance. According to Ricoeur's dialectic, happiness should be what human persons strive for but also rely on chance to take its course. The above understanding implies that happiness is not reductive but holistic, encompassing both happiness and unhappiness.

Nonetheless, in this work, we aim to examine an understanding of happiness that clearly shows the roles others play in contributing to it. The objective here is to inquire if humanoids can be legitimate contributors to our happiness and be placed in the category of what we conceive as the "other". To do the above, we turn to Ricoeur's threads of happiness to see the significance of the "other" in contributing to our happiness.

The first thread centres on the individual as an acting agent who has a personal wish for happiness and seeks fulfilment through personal agency. Yet, human beings exist within complex social systems that shape and constrain their pursuit of happiness. Ricoeur [6] similarly cautions that happiness cannot easily be secured through individual effort alone; such a view, as we also allude to, is limited and ultimately unsatisfying. Additionally, as we will show in the last section, and as presented in the current section, Ricoeur conceives individuals as also beings in community.

The second thread concerns happiness found among those near to us. Ricoeur turns to Aristotle's account of *philia*, particularly friendship grounded in goodness rather than utility. Here, happiness is no longer a private aspiration but emerges through reciprocity—through giving and receiving. Its fragility lies in its shared character. Friendship dispels loneliness and deepens fulfilment, even under the shadow of mortality. As Ricoeur notes, "the death that will one day separate us is not limited to the event that has not yet happened; its shadow stretches ahead of time in the shape of unparalleled fear" [6]. This dimension of happiness resonates empirically in the Harvard Study referenced earlier. Waldinger and

Schulz [1] likewise argue that close relationships are central to healthy ageing and the good life, but they remain fragile, as Ricoeur points out.

However, as we shall argue, friendship (and close relationships such as spouses and family) today increasingly involves human–technology relations, given the proliferation of disruptive technologies. This shift raises questions about how we conceptualise the "other". We return to this point in the next section of our discussion on human-technology relations.

Ricoeur's third thread concerns justice and the "faceless other". A good life, he argues, involves the desire to live within just institutions [6]. Friendship must extend beyond intimate circles to anonymous others within the broader social order. Yet Ricoeur warns against equating justice with the automatic fulfilment of happiness. While justice may serve the common good, happiness must not be imposed through political power or reduced to a public programme [6]. It cannot be mandated by sovereign authority, nor should individuals presume to manage others' happiness.

What is central for our purposes is that the second and third threads both depend upon the presence of others[2]. In friendship, those near to us contribute to our happiness; in justice, our acts towards distant others, whether through peacebuilding, works of charity (including providing food to the homeless individual on our way from work), alleviating poverty, or mitigating climate change, all contribute to shaping our fulfilment. In both cases, happiness presupposes relationality.

Traditionally, these "others" have been human subjects. However, emerging scholarship on human–technology relations challenges this assumption and compels a reconsideration of what counts as the other. For instance, as we show in the next section, some individuals prefer to turn their attention towards non-human tools to derive satisfaction. We ask whether such a turn can be justified as a form of happiness, and whether these non-human entities (AI and social robots) that they rely on for their happiness can legitimately be considered as "others" in the Ricoeurian sense. We now turn to this developing discourse. First, we show why this discourse is important and necessary. Afterwards, we interrogate the notion of otherness emerging from this discourse through Ricoeur's lens.

4. Disruptive Technologies: Blurring the Lines of the 'Other'

Contemporary technology designs and the entire technology ecosystem raise pressing concerns. One such concern relates to the influence of Big Tech on human relationships and societal values. Companies often prioritise economic gain over individual well-being, shaping the logic and norms of communities through their designs [18]. Technologies such as Replika, created as a human companion, occupy the role of the "other" in the lives of some individuals. For example, the very design of Replika is intended to foster companionship and friendship among individuals who feel lonely. Understanding the implications of this requires a conceptual framework, which we offer through the lens of socially disruptive technologies.

Disruptive technologies are traditionally defined in business as those transforming markets [19]. However, the aforementioned conception of disruptive technologies seems too narrow to capture the sense in which it is intended here in the current research. This is because the aforementioned captures only the disruptions on the business side of technology designs. As a result, we adopt an expansive conception of technological disruption in this paper. We adopt a societal perspective informed by the Philosophy of Technology (cf. Hopster [20–22]), which treats disruptive technologies as a concept that encapsulates the socially disruptive nature of technology across all social spheres. We briefly explain the above concept below.

Socially disruptive technologies are designs that distort our traditional social norms, relations, and the way we see the world [20]. One reason these technologies are considered

disruptive is that they are unpredictable and continually challenge our worldviews. Their unpredictability produces novel moral situations that put a strain on moral concepts that are traditional to us. For example, humanoid social robots, designed to resemble and behave like humans, illustrate this phenomenon. Their human-like qualities challenge our conception of the "other," requiring us to adapt our ethical frameworks. This is because human moral patients are no longer the exclusive focus of ethical concern. While this shift in the understanding of moral patiency is taking shape, we present a more sceptical view in the final section of this paper. We interrogate the "moral patiency" of non-human entities, such as social robots and AI. We posit that the aforementioned tools cannot be considered as "moral patients" in any sense.

We understand that historically, technologies were tools serving human ends, and they were neither moral agents nor patients [23]. However, this view is now rapidly changing as social robots are beginning to engage in interactions once reserved for humans (e.g., self-driving cars, diagnostic tools, and socialisation technologies), prompting us to reconsider our established values, norms, and beliefs. Socially disruptive technologies challenge foundational concepts such as personhood, agency, and moral or legal responsibility and reshape our understanding of the "other".

AI, a socially disruptive technology, now performs tasks previously considered uniquely human. For example, historically, communication and writing were reserved for human domains [24], but technologies such as ChatGPT blur these lines, requiring the inclusion of non-human artefacts among those capable of communicating, writing, and producing "new" ideas. This is also similar to the conversation we are having here. AI challenges traditional distinctions such as responsibility, agency, and personhood while giving the illusion of expanding the role of the "other" in promoting human happiness. This illusion then projects the impression that the concept of agency is no longer ascribed exclusively to human persons. Given the role technologies like AI play, and the agential illusions which it projects, it is imperative that we either rethink the frameworks through which we underscore the notion of agency, or interrogate heads-on whether AI can be classified as an agent in the traditional sense which we underscore the concept (in this paper, we are strictly restricted to the capacity to contribute toward human happiness/flourishing as the "other"). However, we think that it is imperative that we rethink our frameworks. This is because there are domains of our societies that must rethink their frameworks; one such area is our legal systems.

The theorist, Turner [25], observes that current laws are insufficient to address AI and social robots, such as automated vehicle drivers. It is imperative that we update and/or amend aspects of the legal framework and laws to accommodate AI. For instance, drivers of cars or other automobiles are no longer only human beings, given the rapid growth of automated vehicles. So, laws regulating driving must change to accommodate these new 'drivers.' The speed fines, which apply to traffic violations, must now apply to social robot drivers. This is like the case of happiness, which this paper explores. Those who contribute to and promote our happiness are no longer only humans; they are also social robots. However, should these social robots be considered as the "other" in the sense through which Ricoeur spells out? We will get to this question in the next section. For now, we show recent entries on human–robot relationships to see whether the role these social robots and AI play justifies the assumption that they could be classified as the "other".

Scholarship on human–robot relations highlights these changes, with individuals forming emotional, romantic, or sexual relationships with robots [26]. Cases such as Kondo, spelt out in the introduction, exemplify the human desire for companionship beyond traditional human–human interaction. For instance, the need for human–human attractions through friendships, dating and marriage is becoming less. Humans are now forging

relationships and friendships with social robots. As Itasu [9] outlines, some people in Japan now prefer interacting with social robots rather than with humans. For instance, individuals like Kondo are marrying robots, and over 1.5 million people prefer not to interact with other humans. Additionally, studies by Emont [7] and Itatsu [9] note that, given the rising culture of social anxiety, people are now more willing to allow humanoid or mechanical robots to assist with their physical exercise.

In care facilities, Emont showed that elderly people preferred receiving care from their care robots rather than from foreign nurses. Additionally, Suzuki et al.'s [8] study showed that individuals with social anxiety preferred that their mechanical robots not be designed to resemble humans. Another study found that people who are fictosexual are not two-dimensional in their relationships: that is, they do not consider the fictional character a substitute for human–human relationships [9]. They prioritise interpersonal-oriented sexuality towards their character, and they warn that having two-dimensional relationships somewhat results in stereotypes and marginalisation, resulting in the invisibility of fictosexual individuals. In their view, such an approach is wrong at best. However, can we consider these relationships as legitimate aspects of relationships that warrant that we conceptualise those to whom these humans fall "in love" with the "other"? It is imperative that we are sceptical about these changes for the following reason.

On the one hand, we acknowledge that the cases and examples mentioned above must be taken into account in assessing the contribution of the aforementioned technologies to human flourishing and happiness. There are good reasons to believe that these technologies are helping the individuals mentioned above flourish in whatever capacities they have employed them. The above point goes without saying for any reasonable interrogator. However, we must also be critical of the above position.

These individuals entering into conjugal relationships with the above technologies may be in a state of self-deception. They may assume they are happy and flourishing based on reports and surveys. These reports may be based on a particular research study's happiness benchmark. In hindsight, these individuals may be experiencing something that could be correlated with a manic episode or Stockholm syndrome—falling in love with an unrecognised corporation responsible for designing these robots. Our concern here is that we ought to be sceptical about some of these reports and surveys on happiness, even though theorists like John Danaher may caution us against such scepticism.

In his work, Danaher [27] warns that we must take caution when we label new sexual identities to avoid pathologising and othering that which should be considered a "part of the ordinary range of human sexual desires" [27]. The impression here is that we must not be quick to consider relationships that are not traditional human–human relationships as problematic. Perhaps the view from Danaher is that, because we do not understand the dynamics of such relationships, it does not invalidate them or those who engage in them. Furthermore, theorists such as Aoki and Kimura [10] claim that social robots have an affective impact that can address vulnerabilities. For example, using the principle of homeostasis, individuals who marry fictional characters, such as Kondo, may be a means of escaping a precarious emotional state, such as suffering and loneliness. However, this is precisely why we raised our scepticism above. If the rationale for such relationships is to circumvent loneliness or some precarious emotional state, then it is imperative to engage directly with the challenges these individuals face, rather than create frameworks that may become obsolete. Nonetheless, we could also say that, in the meantime, the technologies could be used to assist them until they address the main issues. In hindsight, the assistance of these technologies in the above regard may be considered as a contribution towards the flourishing and happiness of these individuals.

Alternatively, as Aoki and Kimura [10] suggest, and a view with which we are more sympathetic towards, we ought to develop new epistemologies surrounding digital intimacy or friendship, besides the dominant need to fetishise and sexualise such relationships. Additionally, if we develop new epistemologies surrounding digital friendships, we may be able to determine whether these technologies promote our happiness to the same degree as the human other in Ricoeur and the Harvard studies, or to any degree at all. However, we think Ricoeur will be sceptical about such a comparison. We hold that Ricoeur may gravitate towards human–human relationships and friendships as more legitimate. Nonetheless, we cannot completely dismiss the idea that these technologies contribute, in some sense, towards human flourishing and happiness, regardless of Ricoeur's possible scepticism.

The point here is that, as it stands, we cannot dispute that these technologies have disrupted traditional concepts such as friendships and relationships. They have also disrupted what it means to contribute towards human well-being and flourishing. Undoubtedly, 'the genie is out of the bottle', and technologies can be profoundly disruptive and/or helpful, but we must remain sceptical about their impacts at all times. AI and social robots come with different challenges. These challenges include sycophancy, arrested development, potential for abuse by authoritarian corporations or regimes, and manipulation by market structures and capitalist tricks. It is imperative that we do not lose sight of these challenges while we engage with the benefits of these technologies. Thus, the justification for developing new frameworks and epistemologies to capture these disruptions and clearly emphasise the roles these tools play in contributing towards our happiness and flourishing is evident.

Some benefits of these tools for human flourishing and wellbeing should not be dismissed. For example, there is positive literature on the ethics of robots [28,29] indicating that social robots will play a unique yet significant role in education, therapy, and hospitals, offering meaningful support. One study (cf. Darling [28]) found that those with autism who have been significantly unresponsive to human therapists have had good interactions with robots. Huggable, a social robot, has shown significant engagement with autistic children in playful interactions [18]. However, we must be cautious about the kinds of developments Darling presents. One might argue that autism is, in part, a disorder of social relating. In their work, Baron-Cohen et al. [30] describe autism as "mind-blindness". This entails difficulty in recognising others' intentions, interpreting social cues, and experiencing emotions vicariously. Against this backdrop, the claim that robots relate more effectively to autistic individuals than humans appears paradoxical. If those who already struggle with human interaction are said to engage better with these technologies, does this not risk confirming, rather than addressing, the underlying relational challenge? In this sense, the argument risks collapsing into a form of *reductio ad absurdum*.

The more pressing question, then, is whether we are willing to undertake the demanding work of relating to one another in all our limitations, which remains essential for sustaining meaningful human relationships. We agree with theorists like Darling [28] that social robots and the relationships which humans create with them can be very enriching; however, they should not be theorised as a replacement for human–human interactions and relationships. The latter part of the prior sentence is what informs our position, on the other hand, that these technologies should not be theorised as the "other" in the human sense.

Additionally, to not lose sight of the primary aim of our argument, we remind the reader that our contention here is not whether AI and social robots can replace humans, but to what extent we can equate them as the "other" in the way we conceive humans. Our argument is clear: as we will show in the next section, these technologies do not satisfy the criteria for otherness in the human sense. As we have acknowledged in the introduction,

these tools can be virtual, partial, or quasi others, but not otherness in the sense that we are interested in this paper—the human other.

For example, in the study on religious social robots, Buddhist robots like Mindar and Xian'er have been instrumental in delineating the role of religious robots with the community of believers. For instance, the robot Mindar is designed to represent the Bodhisattva of Compassion, who is prayed to and bowed to. Xian'er is designed to represent the young "average Buddhist monastic practitioner," with comic characteristics that enable it to be immersed within the community of believers. While Xian'er is a popular social religious robot, it is not prayed to or bowed to [31]. Xian'er has a very clumsy and cute appearance, a sense of humour, and the ability to tell funny stories, and some autonomy to refuse to listen to instructions. Xian'er is also invited to ceremonies and occasions such as cartoon festivals. In Christianity, we also find a number of religious chatbots and AI Jesuses such as AI Jesus, Virtual Jesus, Jesus AI, Text with Jesus, and Ask Jesus [32]. The above instances provide evidence that these bots can be considered partial or quasi-others, but not human others.

Given the role of religious robots within the community of believers, Puzio [33] proposes that these robots be designed in a non-hierarchical manner to facilitate their integration. Additionally, Ahmed [34] has proposed that social robots deployed within religious communities can assume roles such as "religious companions", "counselling agents", "pastor bots", and "psychological chatbots." Again, should they become companions and counselling agents, given their role, our question remains: can we and should we conceive them as the "other" in the sense that is spelt out in Ricoeur? Alternatively, at this point, we contend that this question requires further investigation through empirical research. We also ask the reader to make up their mind about what appeals to their critical thinking. However, in the next section, we present the view we think Ricoeur will be sympathetic to, and our view based on our understanding of Ricoeur.

5. AI and Social Robots as the "Other": A Category Mistake

In the previous section, we showed that the conception of what it means to be the other has been blurred through technological disruption. With the introduction of humanoid social robots, activities once performed by humans to promote happiness and well-being are now carried out by technology. In our exposition of human-technology interactions, we have noted that some people now prefer interacting with social robots rather than with other people. On the one hand, we argue that there are valid reasons to believe that these technologies contribute, in some sense, to human flourishing and happiness, even though we still have to be critical of such reasons. Given where we started, the view that we need other people to promote and contribute to our happiness using the works of Ricoeur and the Harvard Study, it becomes problematic to conceptualise the "other" that promotes/contributes to human happiness and flourishing only as a human person; entities that contribute towards human flourishing and happiness now seem to include technological artefacts like AI and social robots. What are the implications of this shift? Or, is it really the case that the aforementioned technologies can now be described as the "other" in a similar way as the human other? In this section, we seek to clarify the latter in detail, despite hinting at the position which we have taken several times in this paper.

While we also open the above questions to the reader for reflection, we set out our position in this paper. We argue that AI and social robots cannot be regarded as the human "other" in the sense in which Ricoeur theorises "otherness" in relation to human flourishing and wellbeing. Put differently, Ricoeur would be sceptical about labelling AI and social robots as the "other" in a human sense, despite their contributions to human happiness

and flourishing. We briefly outline the grounds for this scepticism before presenting four arguments in support of our position.

Drawing on Ricoeur's account of happiness and otherness, discussed in *Section 3*, we infer that he would maintain that only human beings possess the capacity to be regarded as the "other". Ricoeur may assert that the concept of the "other" has some nuances; yet, human beings occupy a distinctive place and play a distinctive role in relation to otherness. Moreover, only humans understand what happiness and flourishing mean for them and therefore possess the agency to extend or reproduce such meanings in relation to others. By contrast, AI and social robots do not understand what it is to be happy; they merely simulate "happiness". As we show below, such a simulation does not justify classifying them as the "other". However, Ricoeur may admit that these technologies can be considered virtual or partial others, and that their contribution to human flourishing and happiness should be taken into account. Nonetheless, we considered the latter in the previous section; we do not proceed to defend the former. Rather, we take an independent stance, based on our reading of Ricoeur, to show why AI and social robots ought not to be regarded as the "other" in the human sense, regardless of the affinity humans have towards these technologies. We now turn to four arguments demonstrating that it is a category mistake to conceptualise AI and social robots in this way—the "other".

However, before presenting our arguments, we briefly show what a category mistake means. Our understanding of a category mistake stems from Gilbert Ryle [35]. Ryle claims that a category mistake means placing an entity into an ontological category when, in fact, it belongs to a different one. As we will show in the four arguments below, treating or considering AI and social robots as the "other" in a human sense (precisely the Ricoeurian sense) commits a category mistake.

5.1. Argument from Simulation

Our first argument is that AI and social robots cannot be the "other" in the Ricoeurian sense, as they fall between simulation and reality. AI and social robots are meant to stimulate emotions that are not genuine. For example, these robots can feel and behave as if they have human emotions, showing certain facial expressions, uttering certain words, speaking in a humanlike form, and exhibiting humanlike reactions to touch [27]. Through these emotions displayed by their designed features, one might begin to respond to them, in part, in an interactive social way. People will respond to social robots differently based on their emotional features, unlike how they respond to other robots or people. Nonetheless, we must understand that social robots are designed to specifically trigger and stimulate such emotions and behaviours in humans. These effects are by design; there is nothing real about these emotions; they are merely simulations. Thus, they should not be considered as contributors to our happiness in the same way as a human other.

Undoubtedly, the simulations created by social robots affect human experience. For example, Nyholm [36] and Friedman [37] identify three dimensions of humanoid social robots: appearance, behaviour, and maximal human likeness. Social robots like Sophia exemplify this maximal human likeness, combining humanlike appearance and behaviour to elicit anthropomorphic responses. Furthermore, empirical studies show that even partial simulations can enhance human well-being, as users often form meaningful attachments to these robots [38].

One might argue that the key point is not whether robots "truly feel," but how human beings respond to them. Humans naturally attribute humanlike qualities to entities that resemble them in behaviour or appearance. These interactions can produce genuine psychological and emotional benefits, contributing to well-being in ways comparable to human–human interaction in some contexts. While these human responses might have

some normative relevance, we think that social robots may matter only if they enhance human flourishing, and that their simulated emotions may matter ethically, even if the robots themselves do not "feel." However, that is all that they can do. Regardless of the extent to which the simulations and the ethical implications of the emotions which they simulate, we maintain the position that they should not be considered as the "other" in the sense in which Ricoeur uses it, as established in the sense which we spell out in *Section 3* of this current paper.

5.2. Argument from Market Hype

Recently, there has been an overhype and overselling of AI and social robots by the business and AI companies. Emily Bender and Alex Hanna [39] make this point clear in their works by underscoring that AI companies tend to idealise new technologies and use marketing hyperbole to attract sales. We suspect that companion robots and social AI are being overly hyped to gain popularity and increase sales. These market tactics are made possible due to the human tendency to search for a "perfect" partner. However, this idealisation of AI and social robots begins to unravel when one recognises that, in the absence of a physical body, practical competencies, or embeddedness within a shared lifeworld, the AI humanoid offers only a limited form of engagement. What remains is a set of interactions that may be superficially satisfying but lack the depth, reciprocity, and situatedness characteristic of human relationships. At best, such systems provide entertainment through witty exchanges, as seen in large language models like ChatGPT; at worst, they risk fostering a quasi-pathological detachment from the realities of our lived social contexts and/or existence.

While for some individuals, the outcome that they get from these systems may appear "good enough", it is important to note that their judgment may rest on certain assumptions: that the system's responses are consistently "on target", effective, and engaging. However, it is not always guaranteed. We are not oblivious to the fact that such uncertainties are also found in human–human relationships. In ordinary human contexts, sustaining meaningful interpersonal relationships is complex and demanding. The apparent ease with which individuals—particularly those with limited interpersonal skills—may turn to such systems as an "easy alternative" is therefore noteworthy. AI systems can be designed to be accommodating and agreeable in ways that human beings, with their imperfections and constraints, often cannot. Yet it is precisely these challenges, uncertainties, tensions, unpredictability, and moral depth in human interaction that underpin genuine relationality, well-being, and flourishing. The point is that, regardless of how AI and social robots are oversold, they cannot replace the human other, or even be like the human other, given the dynamics of human–human relationships in relation to human flourishing and wellbeing. At best, they can only attain the status of virtual or quasi "otherhood".

5.3. Argument from Moral Patiency

Our third argument is that AI and social robots cannot be the "other" in the sense Ricoeur establishes, because these technologies cannot satisfy the criteria for moral agency and patiency. Drawing from the works of Sparrow [40], we claim that our treatment of social robots has no impact on the robot's point of view, given that they have no mind of their own. In this sense, robots do not feel pain, pleasure, or any other form of suffering, and our actions towards them cannot cause them suffering or happiness. On the contrary, when we think of the other in the Ricoeurian sense, we mean an entity that can feel pain and pleasure, happiness and sadness, suffering and pain. When we do an act of justice, we envisage that those to whom we show care will feel the above feelings in some sense. For Ricoeur, happiness is not only found or received in friendship but also experienced

through the happiness of others. So, the question is whether this last type of happiness can be experienced if robots do not feel suffering or happiness. We highly doubt that AI and robots can experience the above.

While we may be correct in our assertion that social robots do not experience pain, pleasure, or suffering and are not moral patients in the traditional sense [40], we do not advocate treating them as ethically irrelevant. Our rationale here is that human behaviour toward robots reflects moral attitudes that extend to social norms and ethical character. Treating robots violently or degradingly can symbolise and reinforce morally problematic behaviours in humans.

Moreover, as this paper shows, people treat social robots as relational "others," forming friendships, attachments, or even marital-like relationships. Although we maintain the position that the concept of "otherness" ought not to be attributed to these technologies, we also recognise that the relationships and interactions humans have with these systems are genuine human experiences, producing genuine emotional effects. From a Ricoeurian perspective, ethical engagement must consider how actions shape human moral life. The fact that robots are not sentient or the "other" does not negate the ethical significance of these relationships; instead, it highlights the need for new moral frameworks that account for technologically mediated forms of relationality.

5.4. Argument from Human Subjective Experiences

Our final argument draws from Ricoeur's *Time and Narrative* [41]. In the aforementioned, Ricoeur argues that human beings make sense of their lives through the narratives around which they organise their experiences. These narratives enable us to grasp the temporality of our lives and provide the structures through which we interpret our existence. In turn, such narrative structures underpin identity formation, as individuals construct a sense of self through the stories they tell about themselves. However, Ricoeur emphasises that these narratives are not wholly self-created; rather, they are embedded within a pre-existing world of meanings and narratives. While individuals draw on these shared narrative resources, they nonetheless retain the capacity to shape meanings that are, to some extent, distinct, even as they remain situated within a broader lifeworld. By contrast, AI and social robots lack the capacity to create narratives grounded in their lived experience; to think otherwise is a category mistake. AI and social robots do not have the ontological capacity for subjective experiences. To be more direct, AI and social robots lack narrative identity, as they have no lived history from which to constitute a self. While developers may attempt to simulate identity, this identity remains an imposed narrative rather than one grounded in experience.

Through self-narratives from our experiences, human beings can interpret what happiness and flourishing mean to them. This capacity also grounds human ability to relate to others as companions, to recognise challenges, and to respond empathetically. AI and social robots, lacking such capacities, cannot meaningfully occupy the role of the "other". To treat them as such would therefore constitute a category mistake, particularly in the context of their contribution to human happiness and flourishing. Moreover, since these systems cannot experience happiness themselves, they lack any understanding of what it means to intentionally promote it in others.

This is not to deny that AI can contribute to the conditions under which human flourishing becomes possible, as clearly spelt out in the previous sections. For instance, AI technologies can support the development of just institutions and social structures by bringing distant experiences closer—such as enabling access to the stories of those in conflict zones or facilitating technological interventions (e.g., prosthetics, vision, or hearing support) that enhance quality of life. Such contributions may align with Ricoeur's broader

account of the institutional dimensions of happiness. However, they do not entail that AI or social robots fulfil the relational role of the "other".

6. Conclusions

In this paper, we have demonstrated that Ricoeur's framework on happiness is fundamentally relational: it emerges through intimate bonds and just engagement with broader communities. However, we argued that the rise of human–technology relationships, particularly with humanoid social robots, does not negate the relational nature of happiness but rather complicates its application. Social robots, though lacking sentience, can elicit meaningful emotional and psychological responses, enhancing human well-being and happiness in ways that resemble traditional human interactions.

Philosophically, this challenges the strict anthropocentric conception of the "other" and requires expanding ethical frameworks to account for technologically mediated forms of relationality. Ethical engagement must therefore consider how interactions with robots shape human moral attitudes, social norms, and capacities for empathy, reciprocity, and justice, aspects which are necessary for our happiness in the Ricoeurian sense. However, we have clearly argued that AI and social robots cannot be regarded as the "other" in the sense in which Ricoeur theorises "otherness" in relation to human flourishing and wellbeing.

Nonetheless, we recognise that what makes our interactions with humanoids even more lasting than the 'hors lieu' or fragility of human relationships is that these AI bots are not subject to the same mortality as our loved ones. With human beings, the shadow of their mortality always falls on our relationships with them. With AI bots, their 'mortality' is managed through electricity and maintenance. Furthermore, chatbots are always available (24/7), energetic, patient, adaptive, and responsive to our needs. In this regard, they offer much more than our best friends and families for our potential happiness.

While caution is warranted, especially regarding potential harms, ownership, and the displacement of human relationships, the normative significance lies in their impact on human flourishing. The emergence of technologically constituted "others" invites a rethinking of happiness embedded in our sense of relationships. It is imperative that future empirical research assesses whether the conditions under which happiness is cultivated should be prioritised over the ontological status of those who contribute to it. In the current digital era, Ricoeur's insights remain a relevant theoretical framework for future research to carry out the above research. Additionally, the scope must extend to the ethically meaningful roles that technology increasingly assumes in human life.

Author Contributions: Conceptualization, A.H.V. and E.T.U.; methodology, not applicable; software, not applicable; validation, A.H.V. and E.T.U.; formal analysis, A.H.V. and E.T.U.; investigation, A.H.V. and E.T.U.; resources, not applicable; data curation, not applicable; writing—original draft preparation, A.H.V. and E.T.U.; writing—review and editing, A.H.V. and E.T.U.; visualization, not applicable; supervision, not applicable; project administration, not applicable.; funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analysed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Notes

[1] It is imperative to clearly state here that even though Ricoeur speaks of the faceless other, this does not mean a non-human other, but a distant human other. For instance, our act of justice is to ensure that we create favourable environments for human flourishing for those near to others and distant from us.

[2] By the other, we mean the human other.

References

1. Schulz, R.W.M. *The Good Life: Lessons from the World's Longest Scientific Study of Happiness*; Simon & Schuster: New York, NY, USA, 2023.
2. Haybron, D.M. Philosophy and the science of subjective well-being. In *The Science of Subjective Well-Being*; The Guilford Press: New York, NY, USA, 2008; pp. 17–43.
3. Haybron, D. *Happiness: A Very Short Introduction*; Oxford University Press: Oxford, UK, 2013.
4. Haybron, D.M.; Tiberius, V. Well-Being Policy: What Standard of Well-Being? *J. Am. Philos. Assoc.* **2015**, *1*, 712–733. [CrossRef]
5. Ricoeur, P. *Fallible Man*; Kelbey, T.b.C.E., Ed.; Fordham University Press: New York, NY, USA, 1986.
6. Ricoeur, P. *Le Bonheur Hors Lieu, in Où est le Bonheur?* Droit, R.-P.E., Ed.; Le Monde Editions: Paris, France, 1994; pp. 327–341.
7. Emont, J. Japan Prefers Robot Bears to Foreign Nurses. *Foreign Policy* **2017**. Available online: https://foreignpolicy.com/2017/03/01/japan-prefers-robot-bears-to-foreign-nurses/ (accessed on 20 January 2026).
8. Suzuki, T.; Yamada, S.; Nomura, T.; Kanda, T. Do People with High Social Anxiety Prefer Robots as Exercise/Sports Partners? *Jpn. J. Personal.* **2021**, *30*, 42–44. [CrossRef]
9. Itatsu, Y. Social Robots and Regionality: Cross-Cultural Perspectives on Cultural Sustainability. In *Social Robots and Cultural Sustainability*; Springer Nature: Cham, Switzerland, 2026; pp. 103–123.
10. Aoki, B.; Kimura, T. Sexuality and Affection in the Time of Technological Innovation: Artificial Partners in the Japanese Context. *Religions* **2021**, *12*, 296. [CrossRef]
11. Haidt, J. *The Happiness Hypothesis. Putting Ancient Wisdom and Philosophy to the Test of Modern Science*; Arrow Books: London, UK, 2006.
12. Feldman, F. *What is this Thing Called Happiness?* Oxford University Press: Oxford, UK, 2010.
13. Diener, E. Subjective well-being. *Psychol. Bull.* **1984**, *95*, 542–575. [CrossRef] [PubMed]
14. Nussbaum, M.C. *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*; Cambridge University Press: Cambridge, UK, 2001; Volume 97, pp. 543–564.
15. Venter, R. Theology and the (post-)apartheid university: Mapping discourses, interrogating transformation. *Transform. High. Educ.* **2016**, *1*. [CrossRef]
16. Augustine, S. De Beata Vita. In *Augustinus, Over Het Gelukkige Leven*; Ferwerda, R., Ed.; Agora: Baarn, The Netherlands, 1999; pp. 93–110.
17. Verhoef, A. Sisyphus, happiness and transcendence. *South Afr. J. Philos.* **2014**, *33*, 537–546. [CrossRef]
18. Logan, D.E.; Breazeal, C.; Goodwin, M.S.; Jeong, S.; O'Connell, B.; Smith-Freedman, D.; Heathers, J.; Weinstock, P. Social Robots for Hospitalized Children. *Pediatrics* **2019**, *144*, e20181511. [CrossRef] [PubMed]
19. Christensen, C.M. *The Innovator's Dilemma: The Revolutionary Book that Will Change the Way You Do Business*; Harvard Business School Press: Cambridge, MA, USA, 1997.
20. Hopster, J. What are socially disruptive technologies? *Technol. Soc.* **2021**, *67*, 101750. [CrossRef]
21. van de Poel, I.; Frank, L.E.; Hermann, J.; Hopster, J.; Lenzi, D.; Nyholm, S.; Taebi, B.; Ziliotti, E. (Eds.) *Ethics of Socially Disruptive Technologies: An Introduction*; Open Book Publishers: Cambridge, UK, 2023.
22. Friedman, B.; Kahn, P.H., Jr.; Borning, A.; Huldtgren, A. Value sensitive design and information systems. In *Early Engagement and New Technologies: Opening Up the Laboratory*; Springer: Dordrecht, The Netherlands, 2013; pp. 55–95.
23. Nyholm, S. A new control problem? Humanoid robots, artificial intelligence, and the value of control. *AI Ethics* **2022**, *3*, 1229–1239. [CrossRef]
24. Pettit, P. *The Birth of Ethics: Reconstructing the Role and Nature of Morality*; Oxford University Press: New York, NY, USA, 2018.
25. Turner, J. *Robot Rules: Regulating Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2019.
26. Karhulahti, V.-M.; Välisalo, T. Fictosexuality, Fictoromance, and Fictophilia: A Qualitative Study of Love and Desire for Fictional Characters. *Front. Psychol.* **2021**, *11*, 575427. [CrossRef] [PubMed]
27. Danaher, J. Robot Betrayal: A Guide to the Ethics of Robotic Deception. *Ethics Inf. Technol.* **2020**, *22*, 117–128. [CrossRef]
28. Darling, K. *The New Breed: What Our History with Animals Reveals About Our Future with Robots*; Henry Holt and Company: New York, NY, USA, 2021.

29. Elgarf, M.; Calvo-Barajas, N.; Alves-Oliveira, P.; Perugia, G.; Castellano, G.; Peters, C.; Paiva, A. "And then what happens?" Promoting Children's Verbal Creativity Using a Robot. In Proceedings of the 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Sapporo, Japan, 7–10 March 2022.
30. Baron-Cohen, S.; Jolliffe, T.; Mortimore, C.; Robertson, M. Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *J. Child. Psychol. Psychiatry* **1997**, *38*, 813–822. [CrossRef] [PubMed]
31. Travagnin, S. From Online Buddha Halls to Robot-Monks: New Developments in the Long-Term Interaction between Buddhism, Media, and Technology in Contemporary China. *Rev. Relig. Chin. Soc.* **2020**, *7*, 120–148. [CrossRef]
32. Verhoef, A. Artificial Intelligence Jesus Chatbots' Challenge for Theology: An Exploratory Study. *Litnet* **2025**. Available online: https://www.litnet.co.za/artificial-intelligence-jesus-chatbots-challenge-for-theology-an-exploratory-study/ (accessed on 2 March 2026).
33. Puzio, A. The Importance of Relationality in Religious Robots: Why Religious Robots Should Focus on Nurturing Relationships. In *Social Robots and Cultural Sustainability*; Springer Nature: Cham, Switzerland, 2026; pp. 125–151.
34. Ahmed, H.; La, H. Evaluating the Co-dependence and Co-existence between Religion and Robots: Past, Present and Insights on the Future. *Int. J. Soc. Robot.* **2021**, *13*, 219–235. [CrossRef]
35. Ryle, G. IX.—Categories. In *Proceedings of the Aristotelian Society*; Oxford University Press: Oxford, UK, 1938; Volume 38, pp. 189–206.
36. Nyholm, S. Social Robots: Culturally Sustainable or Socially Disruptive? The Case of Humanoid Robots. In *Social Robots and Cultural Sustainability*; Springer Nature: Cham, Switzerland, 2026; pp. 179–203.
37. Friedman, C. Robot Theology: Old Questions through New Media. *J. Ethics Emerg. Technol.* **2022**, *32*, 1–3. [CrossRef]
38. *Intentional Stance Towards Humanoid Robots: Lessons Learned from Studies in Human-Robot Interaction*; Wykowska, A., Ed.; Springer Nature: Cham, Switzerland, 2024.
39. Bender, E.M.; Hanna, A. *The AI Con: How to Fight Big Tech's Hype and Create the Future We Want*; Random House: New York, NY, USA, 2025.
40. Sparrow, R. Virtue and Vice in Our Relationships with Robots: Is There an Asymmetry and How Might it be Explained? *Int. J. Soc. Robot.* **2021**, *13*, 23–29. [CrossRef]
41. Ricoeur, P. *Time and Narrative, Volume 3*; University of Chicago Press: Chicago, IL, USA, 1984; Volume 3.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.