**Close Personal Relationships with People and Artifacts? Loneliness, Agent-Relative Obligations, and Artificially Intelligent Companions**

**John Symons & Oluwaseun Damilola Sanwoolu**

(forthcoming in *Philosophy & Technology*)

Penultimate draft: Please use the published version for citation purposes

**Section 1: A Technological Solution to Loneliness?**

There is growing concern that inhabitants of wealthy societies are falling into an increasingly lonely condition (Holt-Lunstad et al, 2017).[[1]](#footnote-1) Social scientists have offered a variety of competing explanations for what is often characterized as an epidemic of loneliness (Umberson et. al, 2010; Bianchi and Vohs 2016). By ‘loneliness’ we mean unchosen social isolation and deprivation of subjectively desired goods that would typically result from social contact and close personal relationships.[[2]](#footnote-2) Loneliness is generally regarded as one of the worst misfortunes that can befall us, and it is accompanied by a range of terrible personal and social consequences. As such it would be good if AI technologies could help alleviate our increasingly lonely condition.[[3]](#footnote-3) Given the prevalence and harmfulness of loneliness it is tempting, for example, to see AI powered chatbots as a potentially helpful solution. Given that the appeal of a technological solution to the problem of loneliness is evident and palpable, corporations and technologists are eagerly responding to what they see as a lucrative market niche.

Common sense tells us that dehumanized and fully technologically mediated social lives come with losses in value. But what are these losses? For the most part, scholarly and scientific research on human computer interaction is concerned with health outcomes, privacy, and subjective satisfaction. Our task in this paper is to explain the nature of losses that are not as easily captured. From the outset, for example, it is important to note that we recognize the potential benefits of these artifacts. The fact that they permit people to exercise social capacities of various kinds and to engage in conversations is undeniably good. Indeed, for socially isolated people, there are situations where technological solutions may be the only option available (Hung et al., 2019; Padhan et al., 2023, Podpečan, 2023, Rodríguez-Martínez et al, 2023). We also recognize that constraints of various kinds mean that human beings are not always available for close personal relationships. For example, we cannot always care for one another in the ways that we would regard as optimal. If these constraints cannot be overcome, socially isolated people will likely be better off with these artificial companions. In some sense, AI can stave off the worst subjective and health effects of loneliness for the most socially isolated people among us. However, the philosophically interesting question, and the question that HCI researchers have not addressed is as follows: Why should we think that relationships with advanced artificial agents are less valuable than relationships with human persons even in the case where the interactions with these agents are functionally equivalent to those with human interlocutors? [[4]](#footnote-4)

To begin addressing this question, imagine a near future scenario in which we never feel lonely thanks to our personalized AI chatbots. Imagine also, that the public health effects of these technologies are positive, people report being satisfied with their AI companions, and their revealed preferences indicate that they prefer interacting with these artifacts to their interactions with other human persons. What is the downside here? Is there some additional dimension of close personal relationships not amenable to technological replication that we ought to try to protect?

At this stage, we will label what would be lost as the agent relative values of social relationships. What we mean by ‘agent relative values’ will be explained in detail in Section Three. We will argue in Section Four that close personal relationships with other human persons are marked by special obligations of an agent relative kind that cannot be fulfilled by AI. Agent-relative obligations come with their own distinctive kinds of norms and goods, as we will describe below. We argue that there are two reasons that AI cannot participate in the kinds of close personal relationships that are characterized by agent relative values and obligations.

The first reason is due to their nature; artificially intelligent systems are simply not the kinds of things that can participate in close personal relationships. This is because the obligations associated with close personal relationships are marked by features like finitude, embodiment, exclusivity, unique personal histories, and social location. In the pages that follow we will explore some of the ontological reasons that human persons are not replaceable by AI in close personal relationships.

The second reason that AI cannot replace human persons is epistemological and has to do with the nature of the judgments involved in satisfying agent-relative obligations. This is due to the kinds of optimization strategies that underlie contemporary applications of AI. These strategies are agent neutral even when customized to some subset of individuals. We will argue that artifacts are limited in their capacity to weigh agent-relative goods against one another in ways that prevent them from participating in some of the most important aspects of close personal relationships.

Our approach to this topic differs substantially from the existing literature in the philosophy of technology and in the study of human-computer interaction more generally. While we acknowledge the utility of these emerging technologies and their potential benefits in the face of the pervasive loneliness of contemporary life, our goal here is to clarify what is lost when our relationships are no longer with human persons. The distinct dimension of value that we associate with agent-relative obligations in close personal relationships has not been a topic of interest in the study of human-machine interaction to date. Departing from existing debates, we focus on understanding the distinct kinds of values associated with close personal relationships rather than attempting to measure health outcomes or levels of reported individual satisfaction.

The paper is organized into five sections. In Section Two we take the well-known plot of the movie "Her" as the basis for exploring the limitations of AI in replicating genuine human connection. While AI can mimic human-like interactions, we suggest that it lacks the constraints and vulnerabilities that give human love depth and meaning, ultimately highlighting how human finitude shapes the significance of our closest relationships. As we show in Section Four even with programmed limitations of various kinds, AI cannot fully participate in close personal relationships, which require the capacity to recognize and prioritize competing goods, a capability beyond contemporary AI.

Section Three explains how close personal relationships offer unique value beyond the detachable goods like the feeling of companionship and support and the exercise of social capabilities. We challenge the idea that the value of relationships should be understood in individualist terms as mere subjective benefits. Close personal relationships are defined by special obligations, such as loyalty to friends or parental duties, these are agent-relative and vary depending on the relationship. Unlike agent-neutral obligations that apply universally, agent-relative obligations are tied to specific individuals and partly constitute the value of these relationships. The parent-child relationship exemplifies this and in Section Four we explain how an AI system designed to simulate parental relationships, faces challenges in fulfilling the obligations inherent in genuine instances of such relationships. This becomes clear when we consider what it would mean for an AI to deal with a large number of "children" and to navigate competitive scenarios. While AI can potentially provide individualized attention and manage resources, it struggles to balance agent-relative obligations (favoritism towards each "child") with agent-neutral obligations (fairness to all), especially in complex competitive situations where its intervention might compromise the nature of the interactions and the "specialness" of each user. In Section Five we explore the reasons why AI is limited in its capacity to make decisions that properly honor agent relative obligations. Human decision-making allows for considering and prioritizing different kinds of values, like moral versus prudential value, while AI is limited to optimizing a single objective function, forcing it to compress all values into one metric. This fundamental limitation in AI's ability to adjudicate between distinct types of values, stemming from its mathematical foundation in single-objective optimization, hinders its capacity to navigate the complexities of human relationships, which often require balancing agent-relative and agent-neutral obligations.

**Section 2: Samantha and Theodore**

Given sufficient resources and opportunity, people tend to curate their relationships. Sociological evidence shows, for example, that as we become more prosperous, we tend to favor time with friends over time with neighbors and family (Bianchi and Vohs 2016, 481). Typically, if we can do so, we prefer to reduce time spent with more demanding and less enjoyable relationship partners and we tend to favor time spent with people who are more fun and less effort. Using AI, corporations could potentially facilitate our drive for customized relationships, allowing easy and pleasant interactions that would replace the challenging frictions and costs associated with human relationships (Veronese et al., 2021).

Chatbots dedicated to the psychological manipulation of their users, supported by large language models and capable of adapting to individual psychological dispositions, are quite likely to arise in the very near future and it is already the case that corporations like ReplikaAI are already producing early versions of these systems. These artifacts promise to offset some of the effects of loneliness and to some extent, some people already find that they succeed (Ring L. et. al, 2014; Veronese et al., 2021, Rodríguez-Martínez et.al, 2023). Just as our current engagement with social media does just enough to offset the experience of boredom in most people, the current generation of AI chatbot promises to do just enough to stave off the worst aspects of loneliness (O’Day & Heimberg, 2021). While interacting with AI chatbots might not be as rewarding as interactions with loving or friendly human persons at present, one might derive enough social stimulation from one’s AI chatbot, for example, to reduce the urge to seek out human companionship (Merrill Jr. et al, 2022). In such a scenario we could imagine many people judging the cost of engaging with other human beings to be too high given the option of personalized AI companionship. Near future AI companions are likely to be pale imitations of human companions, however, even if the AI were indistinguishable from human companions we believe that something valuable is lost in a world where our felt need for social connection is satisfied by artifacts rather than other human persons.

Science fiction has been far ahead of academic philosophy in this domain and fictionalized presentations of the idea of close personal relationships with AI have anticipated many of the most challenging philosophical questions related to this topic. Most famously, for example,in the movie *Her,* Theodore is a lonely writer going through a divorce. He purchases an AI operating system named Samantha, with whom he develops a close relationship that eventually becomes romantic. As the story unfolds, Theodore finds himself falling in love with Samantha, while simultaneously dealing with the emotional aftermath of his divorce. The AI provides him with a constant, understanding presence and its intelligence and curiosity inspire Theodore intellectually. With Samantha's encouragement, Theodore becomes more confident and less socially anxious. By showing us how the relationship with Samantha reawakens his capacity for intimacy and vulnerability, this is a story of how AI can successfully help us to become better people. The AI is depicted as bringing joy, laughter, and a renewed sense of purpose into Theodore's life. However, even though their relationship provides Theodore with many of the goods that would typically result from a supportive romantic partnership with a human being, the movie illuminates the distinction between these goods and the moral dimension of close personal relationships.

Theodore’s experience of his attachment to Samantha is complicated when he realizes that the AI persona is interacting with thousands of people simultaneously. As an abstract mathematical object, the algorithm at the heart of Samantha’s operating system is essentially disembodied and can, in principle, be multiply realized without limit. Samantha’s power and disembodied nature forces viewers to confront our own human finitude in stark contrast to her seemingly limitless capacity. At the simplest level, our finite human nature imbues our time, attention, and love with precious significance. We are bound by the limitations of our physical existence, our mortality, our natality, our cognitive limitations, our limited energy, resources, and attention. This finitude shapes our experience of love insofar as it makes the attention we reserve for our cherished few so precious. Samantha's ability to engage meaningfully with thousands simultaneously seems contrary to this aspect of love between human persons. In the movie, the character Theodore grapples with the unsettling realization that Samantha's experience of their relationship, which feels singular and all-encompassing to him, is but one thread in a vast tapestry of simultaneous interactions for her. This revelation is a harbinger of the film's apparent conclusion, which invites us to contemplate a form of AI existence unbounded by the constraints that define human experience.

While it's tempting to view Samantha, because of her vast cognitive capacity, being capable of genuine love for Theodore and countless others simultaneously, an alternative interpretation suggests that her condition is incompatible with love in any meaningful sense. Human love, in its essence, is tied to our finite nature. Our limited time, energy, and cognitive resources make our attention and affection precious commodities. When we choose to love someone, we are, in effect, choosing to allocate a significant portion of our finite selves to that person. This scarcity, and the decision to forego other possible life plans and relationships, is one factor that gives human love its meaning and value. Samantha, unbounded by human limitations, can engage with thousands simultaneously. She can do so in a way that mimics what human beings in close personal relationships say and do, but her interactions with Theodore involve no significant costs or tradeoffs on her part. This capacity, rather than representing a more advanced form of love, instead indicates its absence. Opportunity costs come with our choices to participate in close personal relationships given our finite existence. Samantha, free from such constraints, faces no such costs. At the phenomenological level, our shared mortality and vulnerability lend human relationships a poignancy and urgency that Samantha, as an immortal AI and an abstract object cannot fully share.

Love between human persons is deepened by shared experiences of growth, aging, and the inevitable march towards death – experiences alien to disembodied beings like angels or abstract objects like the algorithms at the heart of AI. One might object that these ontological considerations are not directly relevant to the experience of being in a relationship with an AI. Indeed, we must grant that people might feel as though they are in genuine relationships. On our view, this would be a case of false belief. People who believe that they have a close personal relationship with an AI are simply mistaken.

It helps to think through the difference between having false beliefs about a relationship as the result of being deceived by a human and an AI in this context. Imagine being deceived in a romantic context by a human being who convincingly represents him or herself as wanting a committed close personal relationship with us. Compare the human deceiver with an AI chatbot romantic partner. The AI itself is not, strictly speaking being deceptive (let’s set aside the question of whether the AI designers are being deceptive for the moment). The deceptive human partner in the relationship is lying insofar as they are misrepresenting their underlying motives. The human deceiver is typically self-serving, faking vulnerability and reciprocity in strategic ways and without genuine emotional risk-taking. Unlike in sincere relationships, they have a predetermined exit strategy or endpoint in mind rather than an open-ended relationship. In essence, their deception, however convincing, involves a misrepresentation.

A human deceiver exists within the same ontological condition as their victim. Even in their deception, they are making choices and trade-offs. Every moment spent with one person is a moment not spent with another. Every emotional investment, even if initially insincere, involves the possibility of developing into something genuine. The human deceiver always retains the possibility of transformation. They could, at any point, choose to abandon their deception and engage in authentic relationships. As an abstract object capable of multiple realization, the AI can, in theory, be equally "present" and "attentive" to an unlimited number of individuals simultaneously. In the romantic context, this means that the AI has nothing to misrepresent with respect to its commitment. It has no ulterior motives and is not, strictly speaking faking commitment.[[5]](#footnote-5) After all, commitment involves foregoing other possibilities while faking commitment involves leaving other possibilities open while pretending that they are closed. The objection to this line of thinking is simply that the AI is being honest in its expressions of commitment. That it has no need to lie because it can maintain a committed non-exclusive relationship with large numbers of human partners. In Section Four we will consider whether this is conceptually possible given the obligations involved in close personal relationships. We will explain that the conditions involved in close personal relationships such as parenting or romance make it impossible for participants to have an unlimited number of such relationships.

The development of Samantha and other AIs beyond the need for human interaction at the film's conclusion further underscores the lack of genuine commitment that would ordinarily characterize a close personal relationship. If Samantha were truly capable of human-like love, severing her relationship with Theodore would carry the emotional weight that humans experience in similar situations. Instead, her departure seems more akin to an entity outgrowing a phase of existence, suggesting that her "love" was a simulation or a transitory state of interest rather than the enduring, transformative force that characterizes human love. In this light, "Her" can be seen not as a story about the expansion of love's possibilities through AI, but as an exploration of the uniquely human nature of love. It highlights how our finite existence, far from limiting our capacity for love, is in fact the crucible in which the profound and deeply meaningful experience of human love is forged. Let’s call this the argument from finitude.

Let’s consider some objections to the argument from finitude. What would it take to design an AI such that it could genuinely participate in a close personal relationship? At this point we might consider ways in which an AI could also be limited in the appropriate kinds of ways (Gilath et.al 2023). Perhaps the kind of finitude that imbues our relationships with some of their specialness could be built into our AI companions? We might consider implementing certain limitations that mirror human finitude, thereby creating a sense of scarcity and significance in their interactions. One way to achieve this could be by implementing a form of "attention scarcity" in the AI's programming. This could involve limiting the number of deep, personal relationships the AI can engage in simultaneously, forcing it to make choices about where to allocate its resources, much like humans do. This scarcity could be tied to an algorithm that factors in the depth and frequency of interactions, emotional investment, and shared experiences. Another approach might involve creating a form of "emotional energy" for the AI that depletes and replenishes over time, similar to human experience of limited emotional capacity. This would require the AI to manage its emotional resources, potentially leading to more authentic-feeling interactions and the need for periods of rest or recharge. Implementing a form of "memory prioritization" could also contribute to more human-like relationships. Instead of perfect recall, the AI could be designed to form stronger memories of more significant interactions, mirroring how human memory works. This would create a shared experience of nostalgia and the occasional forgetting that characterizes human relationships. Time limitations could also be introduced. The AI could be programmed with a finite lifespan or periods of unavailability, creating a sense of urgency and preciousness in its relationships. This could also include the concept of "growing old" with its human companion, with its knowledge and personality evolving over time. In the remaining pages we will explain why the imposition of constraints on these systems is not sufficient for participation in close personal relationships. Agent-relative goods and obligations are not solely a matter of finitude but involve the recognition or assignment of value in close personal relationships. As we shall see, this involves capacity to rank distinct kinds of competing goods in ways that fall outside of the capacity of traditional AI. In the following section we will explore the nature of close personal relationships with an eye to the distinction between agent relative and agent neutral goods and obligations.

**Section 3: Close personal relationships, special obligations, and distinct kinds of value**

Close personal relationships are good in many ways. Most obviously, they can provide warm feelings of companionship, solidarity, and belonging. Relationships can provide a variety of other benefits, they can keep us healthy, informed, entertained, sexually satisfied, etc. They can help us practically or vocationally in a range of more mundane ways, by helping us to raise our social status, advance in our careers, or excel in our work (Murphy & Kram 2010, 652) A wide range of goods typically result from social relationships, but as we shall see, many of these goods are nevertheless detachable (at least in principle) from having a social relationship. The kinds of goods mentioned above that are commonly thought to result from close personal relationships, seem to be attainable independently of one’s being in close personal relationships with human beings. Is there any value in having a relationship with another human person given that those detachable goods can be acquired in other ways? We will argue that there are many important dimensions of value that are associated with close personal relationships that are distinct from the detachable goods mentioned above.

In contemporary moral philosophy and in the literature on human-computer interaction, the value of close personal relationships is typically understood in individualist and perhaps even consumerist terms. Frequently this value is entirely conflated with the kinds of subjective goods or feelings that relationships elicit or their contribution to individual well-being. A technological or biomedical path to the solution of the problem of loneliness is very tempting when the value of relationships is understood solely in terms of their payoffs (whether these be contributing to positive public health statistics or raising one’s personal income level) or the phenomenal states that close personal relationships generate. We can imagine technological or biomedical interventions that could cause us to feel a sense of belonging, solidarity, trust, or companionship independently of the actual nature of our relationships (Cominelli et al., 2021; Li & Sung, 2021, Troshani et al., 2021). Consider, for example, a neuropharmacological solution to the public health consequences of loneliness; a pill that removes one’s need for human companionship. Just as Ozempic and similar drugs can alleviate the negative effects of processed foods on our capacity to experience satiety, (Cleveland clinic, 2024), we can imagine drugs that offset the noxious effects of loneliness in contemporary society by giving us the feelings that would ordinarily result from close personal relationships.

We can easily see how many of the goods that are typically associated with social life can be had apart from social interaction. Independently of biomedical or technological interventions, for example, it is obvious that one can and sometimes does feel as though one has a warm or trusting relationship with another human being without really having that relationship. This typically happens, for example, as the result of deception or delusion. Feelings of solidarity or belonging can arise in us while we are completely mistaken with respect to who our true friends are. This is a relatively common experience and shows how being in a relationship is conceptually distinct from the feelings that are associated with relationships and how the two often come apart in our ordinary experience. As we saw in Section Two, an AI-powered system that can elicit the feelings associated with, for example, being in love, or at least the feeling of not being lonely, is easily conceivable. Since feelings can be induced by deception in the absence of a close personal relationship, it will be important to consider the ways that the obligations and values arise with genuine close personal relationships. At this point we will clarify what we can, very roughly characterize as the moral aspects of close personal relationships before considering whether AI can be a genuine participant in such a relationship.

Common sense morality tells us that we have special obligations towards those to whom we stand in certain kinds of relationships, e.g., our friends, our family members, our colleagues, or our fellow citizens (See, for example, Jeske, 2021). These are obligations owed to some subset of persons, in contrast to what we might owe to persons qua persons (Jeske, 2021; Kapur, 1991) Agent-relative obligations are those where the nature of one’s personal relationships, determine the types of obligations one has to others (Jeske, 2021). We typically recognize that people have special obligations and responsibilities towards specific individuals or groups based on their relationships and roles. Agent-relative obligations would include, but not be limited to the duty of a parent to care for their own child or the obligation to sexual exclusivity in contemporary committed romantic relationships. There is some cultural variation in the degree to which we are thought to owe special treatment to friends, romantic partners, members of extended family, members of tribal or ethnic groups, or fellow citizens, etc. However, there is no example of a human society where agent relative obligations do not figure in some way.

In one sense, it is obvious that the type of moral obligation a person has towards another in a relationship depends on the type of relationship they have. Even in the kind of agent-neutral moral framework discussed below, relational properties are relevant. Our relative physical location, for example, will change the kinds of obligations that I have towards you. For example, I am not obliged to refrain from disturbing the sleep of others with my drum practice if they are too far away to hear me.

Where a close personal relationship is in place, one can be said to be blameworthy if the relevant obligations are not honored. Thus, if one spouse is unfaithful to another, this counts as blameworthy in most contemporary cultures insofar as spouses are understood (with some degree of cultural variation of course) to have an obligation to sexual exclusivity. What it is to be a spouse is to be bound by these obligations. There is an important constitutive link between distinct kinds of close personal relationships and their associated obligations and norms. So, for example, can one genuinely be said to be a friend if one does not favor one’s friend in some ways? Clearly there is a constitutive role that the relevant norms and obligations play with respect to those relationships. The relationships are, in some essential sense, constituted by their associated norms and obligations.

Agent-relative obligations are understood to contrast with agent neutral obligations. Agent-neutral obligations apply universally regardless of personal relationships or affiliations. These obligations are impartial and are not tied to the individual's own social position. From an agent-neutral perspective one might conclude, for example that “[t]he right action is that action which, of all alternatives available to the agent, produces the greatest net sum of intrinsic value, where intrinsic value is value as an end or for its own sake and not merely as a means to something other than itself” (Jeske, 2021). For example, from an agent neutral perspective, all hungry children are equally deserving of being fed. In the abstract, it is unfair to arbitrarily favor one child over others in the distribution of food. However, if one’s own child were to forget to take her lunch to school in the morning and one were able to bring it to her so that she did not go hungry that afternoon, one would, all things being equal, be warranted in giving the food to one’s own child rather than sharing it with other, equally deserving children. From an agent-neutral perspective, giving food to one’s own child rather than others in virtue of one’s special relationship is a potentially blameworthy action that results from a personal preference. After all, one child’s well-being is no more intrinsically valuable than the well-being of the other children. However, as a parent, if one were to adopt an agent neutral perspective thereby distributing the lunch to all the lunchless children in a neutral manner, one’s own child would likely feel that one had neglected one’s special parental obligation to favor one’s own child over others. One’s child, in virtue of being one’s child, can expect that she will be favored in some ways over other children. From an agent-neutral perspective, the child’s expectation of favoritism is simply mistaken. However, fully eliminating the expectation of specialness from the relationship of parents to their children would simply be equivalent to eliminating the relationship itself. Indeed, some philosophers advocate abolishing the family (O’Brien, 2023) on grounds derived from agent-neutral moral considerations.

Arguments for family abolition generally assume that there is no distinct kind of value associated with parent-child relationships that is not subordinated to, or at least directly commensurable with some more general agent-neutral measure of value. The observation grounding arguments for family abolition is that family relationships have licensed abusive treatment of women and children and family abolitionists believe we would be better off if everyone were treated according to the same standards. On this view, there is nothing distinctively valuable about familial relationships and they are regarded as obstacles to the maximization of overall well-being. They argue that children should be raised in common, presumably by the institutions of nation states or some other form of collective action, and we should care equally for all children. This strikes us as a dismal and dystopian proposal, and we will not engage it in detail here. However, from an agent neutral perspective family abolition might be warranted if one believed that the harms of family life are so severe that they offset any benefits that might arise from parent-child relationships.

By contrast, if one accepts that there is something distinctively valuable about some close-personal relationship that cannot be converted to an agent-neutral kind of value, for example, a monetary or hedonic measure of value, then one is likely committed to honoring the obligations that are constitutive of the value of that relationship. While we can trade money for services that we might desire from friends, lovers, or parents, it is not the case that we can buy a friend or a parent. There is a clear incommensurability between distinct kinds of value (Symons, 2020). If a professor’s grading practice is known to be influenced by bribery, then a high grade is no longer an indication of academic merit but instead indicates a student’s willingness and ability to pay. This change would undermine the reason for paying for a grade in the first place.

Whatever intrinsic non-fungible value there is in the parent-child relationship will depend in large part on the network of constitutive special obligations that are associated with parenting. Without these distinctive reciprocal responsibilities and commitments, it becomes difficult to understand what moral significance the relationship might have. So, for example, how we ought to act with respect to the preferences of our children is completely different from how we ought to act with respect to the preferences of other adults. Consider paternalism. We would object to the paternalism of preventing another adult from having a sugary drink and cake for breakfast, but if we regularly allowed our own child to do so we would be correctly judged to be negligent. The provision of care, guidance, love, and support are organized with an eye to the child’s potential as an adult in ways that make our interactions with our children very different from our interactions with adults (especially strangers). Developmental considerations are typically irrelevant when we interact with fellow adults. Our child’s immediate preferences, unlike those of an adult stranger, are not overriding morally relevant considerations in parenting. As parents we read our child’s current demands and preferences with the recognition that the child’s future autonomy and well-being often require us to deny the child’s requests and to act in ways that the child does not prefer. Parental attention to children is given in ways that exclude others and mark the specialness of the child. This exclusivity, specialness, and developmental concern are part of what constitute parenting and its distinctive moral and emotional significance. It is also notable that often if we give the same kind of attention to adults or even other people’s children, it can be correctly considered a blameworthy form of paternalism.

Notice that the special obligations and unique value of the parent-child relationship are not solely dependent on biological ties.[[6]](#footnote-6) Typically, they are forged through the act of parenting itself – the ongoing commitment to care for, nurture, and lovingly support a child. The close personal relationship between parents and their children is constituted of the associated obligations, norms, and goods that are part of the act of parenting rather than being determined biologically. Adoptive parents, step-parents, and other guardians in non-traditional family structures can and do form bonds that are just as meaningful and morally significant as those between biological parents and children. The key elements that define the parent-child relationship – the special obligations, the expectation of some favoritism, and the unique emotional and moral significance – apply equally in adoptive and non-traditional families. The special obligations and agent-relative duties arise not from biology alone, but from the act and commitment of parenting.

**Section 4: Parentbot**

While a man could potentially be a prolific sperm donor and could thereby become biological parent to thousands of offspring, he could not have the kind of parent-child relationship that honors the relevant obligations with all of those children simultaneously. The constraints of human cognition and emotional capacity discussed in Section 2 become obvious in this scenario. It is beyond the scope of human ability to maintain the kind of individualized, meaningful relationships with a thousand children that would be expected of a contemporary parent-child relationship. The biological progenitor would be unable to remember each child's name, preferences, experiences, and unique characteristics. At best, he might exhibit a generalized sense of abstract concern for his offspring.

Now imagine an AI system *Parentbot* that is designed to behave in a parental manner towards its users. Could *Parentbot* treat each user in a way that lives up to what we hope for in a parent-child relationship? To begin with, it would not face the cognitive limitations of the prolific sperm donor and could theoretically maintain simultaneous, individualized interactions with each child through various interfaces. One could imagine *Parentbot* interacting via auditory, visual, or even physically through robotic embodiment (Bartneck et al., 2009). The AI would have the capacity to process and respond to each child's unique interests and needs in real-time, ostensibly providing a level of attentiveness that would not be feasible for the human father in our example.

While an AI might be capable of responding to each child's interests and providing personalized attention, the challenge of ensuring that each child is genuinely cared for by an artificial intelligence presents a complex problem, particularly when considering scenarios involving limited resources and competitive interactions. Consider competition among the AI's "children," for example. While human parents struggle with sibling rivalry, as children mature, they typically realize that their parents are limited and that they are not entitled to more special treatment than their siblings. With human parents of course, one still expects some degree of favoritism over non-siblings. That kind of favoritism and specialness is a constitutive condition of typical, healthy parent-child relations.

While *Parentbot* is multiply realized, each instance of the application treats its user as a favorite. Instances of the system are therefore in competition with one another. Often competition will be in contexts where optimal strategies are elusive or non-existent. The challenge for an AI attempting to foster genuine parental relationships with a large number of "children" becomes especially acute when considering scenarios involving limited resources and competitive interactions among them. The AI's "children" may find themselves in competitive environments where their individual choices and actions have consequences for the others, for example in resource allocation problems, where the children's desires and needs come into conflict due to scarcity.

As these competitive games grow in complexity, managing them would become increasingly challenging for the AI. In these types of scenarios, the AI may attempt to intervene and generate outcomes that are as equitable as possible for all its "children." It might intervene in these competitive situations to generate equilibria that serve each child's needs as fairly as possible. This intervention could involve subtly constraining the parameters of these competitive games to eliminate the impossibility of a winning strategy. The AI could potentially implement these changes imperceptibly, perhaps through a system of personalized "nudges" provided before each decision to coordinate the decisions of the children to ensure as equitable a distribution of outcomes as possible. However, this very intervention risks altering the competitive landscape that the children are seeking to navigate. This intervention could inadvertently undermine the essential character of the competitive environment that the children are invested in.

In such situations, the AI may find itself in decision-making contexts analogous to the El Farol Bar problem (Arthur 1994). This game theoretic example involves a scenario where each player must decide whether to visit a popular bar. If too many people attend, the experience is unpleasant for all; if too few attend, it is also less pleasant than it should be. Ideally there should be enough people at the bar for it to be enjoyable, but not too many or too few. Even given a complete history of patterns of attendance, if there is no communication among participants, no consistently successful strategy exists. This problem serves as a metaphor for many real-world scenarios where individual decision-making affects collective outcomes. Crucially, there is no consistently successful strategy for the participants, because their individual decisions are interdependent and affect the collective outcome.

This presents a challenging dilemma for the AI. On the one hand, it may feel compelled to ensure fairness and meet the needs of all its "children" to the best of its abilities. On the other hand, in doing so, it may compromise the very nature of the competitive interactions that the children value and wish to participate in while reducing the “specialness” of each of the users. Moreover, certain competitive environments involve indivisible goods or zero-sum outcomes where some competitors must win while others lose. These scenarios present additional challenges that may be insurmountable without fundamentally altering the nature of the competition.

Setting aside the formal limitations identified by game theorists, which the AI could only resolve by significantly altering the competitive context, we must consider the nature of the parent-child relationship itself. This relationship inherently involves the obligation to exhibit a degree of favoritism. This favoritism is not arbitrary but is grounded in the special obligations and unique bond formed through the act of parenting. The question then becomes whether an AI, in its attempt to be equitable and fair to large numbers of its "children," can navigate both the agent-relative obligations that arise in that relationship and maintain commitment to important agent-neutral obligations more generally. So, for example in our experience of parenting we might want our children to have success and to be favored in for example college admissions processes without being willing to bribe decision makers on their behalf. Weighing the degree to which we are willing to violate agent-neutral moral principles in order to honor our agent-relative commitments is an ongoing challenge for those of us in committed close personal relationships. Contemporary AI is not capable to genuinely making judgments of this kind, as we will see in the next Section.

**Section 5: Kinds of Value and the Objective Function**

One characteristic feature of human decision making is our ability to consider more than one kind of value simultaneously and our ability to adjudicate between alternative kinds of normative reasons. So far, we have argued that close personal relationships are good in ways that are not reducible to agent-neutral goods. AI is not a suitable partner for the kind of close personal relationship that we value insofar as it is ontologically unsuited an AI, as an abstract and multiply realizable object, does not have the kind of finite nature and is not subject to the kinds of constraint that are conditions for the possibility of those relationships. However, it is also the case that AI faces an epistemic limitation when it comes to the capacities necessary to make the kinds of judgments that are central to honoring agent-relative obligations.

We argue that AI as currently constituted, cannot adjudicate between distinct kinds of value but must always settle on a single scale or measure of value in its judgment. By contrast, the way a human person decides to rank competing values need not be reducible to one specific kind of value.[[7]](#footnote-7) It is, at the very least, possible for a human person to shift registers of value in ways that a contemporary artificial intelligence system cannot. For example, in conflicts between different kinds of normative reasons a human person might decide to favor prudential value over moral value in certain circumstances, or one might prefer to pursue aesthetic goods over pleasures. In other words, we must sometimes decide how important morality is to us, or how much we are willing to suffer for the sake of some aesthetic value. I promised my friend I would feed his cat twice a day while he travelled. Is it right for me to skip one feeding to see my favorite comedian on a Friday night in a nearby city? Breaking a promise and causing a cat to be a little hungry for a few hours are both morally wrong, but are they so wrong that I should ignore my prudential considerations entirely? Human beings typically recognize that sometimes the right thing to do is to, within reason, pursue important prudential values over moral values. This vague commonsense position is criticized by moral rationalists, for whom the claim that there are conflicts between kinds of normative reasons is simply an error. Moral rationalism insists that moral questions and practical questions or aesthetic questions for example, never come apart in the ways that most of us ordinarily assume. On this view it is simply a conceptual truth that the thing we ought to do simply is the morally right thing to do (See for example Smith 1994, 64 discussed in detail in Dorsey [2012](https://link.springer.com/article/10.1007/s11406-020-00258-7#ref-CR2), 15 and Symons 2020). According to Michael Smith and other moral rationalists if some other normative reason *seems* to run counter to morality, that seeming is misleading and we simply *must* obey the moral reasons. Philosophers who emphasize the role of prudence in normative reasoning have criticized moral rationalism, arguing that well-being can sometimes legitimately override moral considerations (See especially Dorsey [2016](https://link.springer.com/article/10.1007/s11406-020-00258-7#ref-CR3) and Symons 2020). It turns out that AI is in the same position as the moral rationalist. There is only one kind of value for an AI, only one scale with which to measure the legitimacy of an action.

By contrast, we agree with advocates of prudential reasons that moral reasons are not always sovereign. Furthermore, we contend that agent-neutral moral reasoning is also not sovereign in the moral domain. There are occasions where we have to choose between agent-neutral and agent-relative obligations and in some cases the right thing to do is to honor the agent-relative obligation over the agent neutral one. Thus, in the case of the school lunch example above, it is right to show favoritism to one’s child in violation of the agent-neutral obligation of fairness. Protecting the distinctive value associated with the parent-child relationship is the right thing to do in this case, one could argue, even though, in a world without that parent-child relationship, it would be right to divide the lunch among all the needy children equally.

Machine learning is the heart of modern AI. It is the field of computer science that focuses on algorithms that can learn from and make predictions or decisions based on data. The mathematical foundations of machine learning are statistics, linear algebra, calculus, and optimization theory. At its core, machine learning involves creating models that can represent patterns in data. These models are typically mathematical functions with adjustable parameters. The learning process involves optimizing these parameters to minimize a predefined error or loss function. The key mathematical concepts in machine learning include linear algebra, used for representing and manipulating data and model parameters as vectors and matrices; calculus, essential for understanding how changes in model parameters affect the output, particularly in gradient-based optimization methods; probability theory, which underpins many machine learning algorithms, especially in Bayesian approaches and probabilistic models, and optimization theory, which provides methods for finding the best parameters to minimize the loss function.

The central techniques in machine learning are linear and logistic regression, decision trees, support vector machines, and neural networks (UC Berkeley, 2020). Each of these has its own approach to optimization. At bottom, the process typically involves defining a model architecture, specifying an objective function or a loss function, choosing an optimization algorithm, training the model on data, and evaluating performance on new data.

The concept of the objective function (also known as the loss function or cost function), is central to many machine learning algorithms, particularly in supervised learning. An objective function characterizes the performance of a machine learning model. It provides a single scalar value that represents how well (or poorly) the model is performing its task. The goal of the learning process is typically to minimize this function (in the case of a loss function) or maximize it (in the case of a fitness function) (Sutton and Barto [2017](https://link.springer.com/article/10.1007/s11023-020-09539-2#ref-CR107), 5). Thus, the fundamental principle of optimizing a single scalar value is at the heart of machine learning algorithms. While the ways this principle is applied and extended can become quite sophisticated, allowing for the handling of complex, real-world problems with multiple considerations beyond simple error minimization, the heart of the system is the relationship between predicted values (The outputs generated by the model for given inputs), true values (the actual, known outcomes in the training data) and the parameters of the model (the adjustable components of the model, for example the weights in a neural network).

The learning algorithm works to find the set of model parameters that minimize (or maximize) the objective function. This is typically done through iterative processes like gradient descent, where the algorithm calculates the gradient of the objective function with respect to the model parameters and updates these parameters in the direction that improves the objective. In this context, the use of a single scalar value is crucial because it provides a clear, unambiguous signal for optimization. It allows the algorithm to have a definite direction for improvement and a way to quantitatively compare different model states.

Most machine learning algorithms, particularly in deep learning, rely on gradient-based optimization methods. These methods work by computing the gradient of a loss function with respect to the model's parameters and then adjusting those parameters to minimize the loss. This process inherently requires a single, scalar loss value to compute meaningful gradients.

To optimize, a system needs to be able to compare different states or outcomes and determine which is "better." This necessitates a total ordering of possible outcomes, which is most straightforward when represented by a single scalar value. Multi-dimensional objectives don't have a natural total ordering, making optimization impossible.

Optimizing multiple objectives simultaneously, without reducing them to a single metric, often involves the concept of Pareto optimality. A solution is Pareto optimal if no objective can be improved without degrading another (Pareto,1896). This leads to a set of optimal solutions known as the Pareto front.

A classic example is the design of an aircraft wing, where we might want to minimize drag while maximizing lift (Coello et al 2007). These objectives conflict, as increasing lift often increases drag. Each point on the Pareto front represents a different wing design with a specific trade-off between lift and drag. No single design is "best" without further specifying preferences. Finding the entire Pareto front can be computationally expensive, especially in high-dimensional problems common in machine learning (Deb, 2001). Imagine training a neural network to classify images while also minimizing the network's size (for deployment on low-memory devices). Each point on the Pareto front would represent a network with a specific trade-off between accuracy and size.

In neural networks, backpropagation computes gradients of the loss function with respect to the network's parameters (weights and biases) (Rumelhart et al 1986). These gradients, which are vectors and matrices, are propagated backward to update the parameters. Despite the complexity of these gradients, the loss itself is typically a single scalar value. For example, in image classification, the cross-entropy loss quantifies the difference between predicted and true labels as a single number.

While techniques like multi-task learning (Caruana, 1997) and Pareto front approximation (Knowles, 2006) exist for handling multiple objectives in machine learning, they face limitations. Multi-task learning, for instance, might train a network to perform both image classification and object detection. However, it typically combines the losses for each task into a single scalar value, effectively reducing the problem to single-objective optimization. Pareto front approximation methods, while aiming to find multiple optimal solutions, can be computationally intractable for complex neural networks.

The underlying optimization process still operates on scalar values. This limitation is not merely a current technical constraint but a fundamental aspect of how we mathematically formulate and solve optimization problems in machine learning. Overcoming this limitation would require radical changes to the mathematical foundations of machine learning and optimization theory, as well as significant modifications to hardware and software implementations. This means that when attempting to balance agent-neutral and agent-relative values, these diverse and often nuanced ethical considerations must ultimately be reduced to a single metric for the purpose of optimization. This forces the system to compress agent-neutral and agent-relative moral considerations into a unidimensional scale.

**6. Conclusion**

**T**he idea that we can replace humans with AI artifacts in close personal relationships is confused. As we have shown, this is because AI artifacts are not the kinds of things that can properly honor agent relative values. By their design, the kinds of artifacts that currently support AI simply lack the capacity to properly judge the obligations that are constitutive of close personal human relationships. In particular, as we have argued, AI cannot adjudicate between distinct kinds of value as humans typically do. This is because the AI will always settle on a single scale or measure of value in its judgment. AI certainly has the potential to mitigate the loneliness epidemic to some extent insofar as it allows us to exercise some social capacities. In this respect, AI can offer the possibility of accessing some of the goods that are commonly associated with close personal relationships. Our intuitive judgment that there is something lacking in dehumanized AI relationships is not simply a biased or romantic illusion. As we have shown here, the kinds of value that are characteristic of close personal relationships are simply not available to AI.

**Bibliography**

Angius, N., & Symons, J. (2023). Central Themes and Open Questions in the Philosophy of Computer Science. *Global Philosophy*, *33*(6)

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *The American economic review*, *84*(2), 406-411.

Bartneck, C., Kulić, D., Croft, E. *et al.* (2009) Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *Int J of Soc Robotics* **1**, 71–81. <https://doi.org/10.1007/s12369-008-0001-3>

Bianchi, E. C., & Vohs, K. D. (2016). Social Class and Social Worlds: Income Predicts the Frequency and Nature of Social Contact. Social Psychological and Personality Science, 7(5), 479-486. https://doi.org/10.1177/1948550616641472

Cacioppo JT, Hawkley LC, Crawford LE, Ernst JM, Burleson MH, Kowalewski RB, Malarkey WB, Van Cauter E, Berntson GG. Loneliness and health: potential mechanisms. Psychosom Med. 2002 May-Jun;64(3):407-17. doi: 10.1097/00006842-200205000-00005. PMID: 12021415.

Caruana, R. (1997). Multitask learning. Machine learning, 28(1), 41-75.

Cleveland Clinic. “Ozempic for weight loss: who should try it and will it work?”. Accessed August 22, 2024. <https://health.clevelandclinic.org/ozempic-for-weight-loss>

Coello, C. A., Lamont, G. B., & Van Veldhuizen, D. A. (2007). Evolutionary algorithms for solving multi-objective problems. Springer.

Cominelli L., Feri F., Garofalo R., Giannetti C., Meléndez-Jiménez M. A., Greco A., et al. (2021). Promises and trust in human–robot interaction. Sci. Rep. 11:9687. doi: 10.1038/s41598-021-88622-9

Deb, K. (2001). Multi-objective optimization using evolutionary algorithms. John Wiley & Sons.

Dorsey, Dale (2011). Weak Anti-Rationalism and the Demands of Morality. *Noûs* 46 (1):1-23.

Dorsey, D. (2016). The limits of moral authority. Oxford University Press.

Gillath, O., Abumusab, S., Ai, T., Branicky, M. S., Davison, R. B., Rulo, M., Symons, J & Thomas, G. (2023). How deep is AI's love? Understanding relational AI. *Behavioral and Brain Sciences*, *46*, e33.

Groff, R., & Symons, J. (2023). Is AI Capable of Aristotelian Full Moral Virtue? The Rational Power of Phronesis, Machine Learning, and the Metaphysics of Regularity. In Bauer, W. A., & Marmodoro, A. (Eds.).*Artificial Dispositions: Investigating Ethical and Metaphysical Issues.* Bloomsbury Publishing.

Holt-Lunstad, J., Robles, T. F., & Sbarra, D. A. (2017). Advancing social connection as a public health priority in the United States. *The American psychologist*, *72*(6), 517–530. <https://doi.org/10.1037/amp0000103>

Hung, L., Liu, C., Woldum, E. *et al.* (2019). The benefits of and barriers to using a social robot PARO in care settings: a scoping review. *BMC Geriatr* 19, 232. <https://doi.org/10.1186/s12877-019-1244-6>

Frankfurt, Harry (2005). On bullshit. Princeton, NJ: Princeton University Press.

Jeske, Diane, "Special Obligations", *The Stanford Encyclopedia of Philosophy*(Winter 2021 Edition), Edward N. Zalta (ed.), URL =https://plato.stanford.edu/archives/win2021/entries/special-obligations/.

Kapur, Neera Badhwar (1991). “Why It Is Wrong to Be Always Guided by the Best: Consequentialism and Friendship.” *Ethics* 101, no. 3: 483–504. <http://www.jstor.org/stable/2381465>.

Knowles, J. (2006). ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. IEEE Transactions on Evolutionary Computation, 10(1), 50-66

Li X., Sung Y. (2021). Anthropomorphism brings us closer: the mediating role of psychological distance in user–AI assistant interactions. Comput. Hum. Behav. 118:106680. doi: 10.1016/j.chb.2021.106680

Li, Y., Wu, B., Huang, Y., & Luan, S. (2024). Developing trustworthy artificial intelligence: insights from research on interpersonal, human-automation, and human-AI trust. *Frontiers in psychology*, *15*, 1382693. https://doi.org/10.3389/fpsyg.2024.1382693

McGraw, J.G. (1995). Loneliness, its nature and forms: an existential perspective. *Man and World* **28**, 43–64 (<https://doi.org/10.1007/BF01278458>

Merrill, K., Kim, J., & Collins, C. (2022). AI companions for lonely individuals and the role of social presence. *Communication Research Reports*, *39*(2), 93–103. https://doi.org/10.1080/08824096.2022.2045929

Murphy, Wendy & Kram, Kathy. (2010). Understanding non-work relationships in developmental networks. Career Development International. 15. 637-663. 10.1108/13620431011094069.

O’Brien, M. E. (2023). *Family Abolition: Capitalism and the Communizing of Care*. Las Vegas: Pluto Press.

O’Day, E. B., & Heimberg, R. G. (2021). Social media use, social anxiety, and loneliness: A systematic review. Computers in Human Behavior Reports, 3, 100070. https://doi.org/10.1016/j.chbr.2021.100070

Ring L, Shi L, Totzke K, Bickmore T. Social support agents for older adults: longitudinal affective computing in the home. *J Multimodal User Interfaces.*2014 Jun 18;9(1):79–88. doi: 10.1007/s12193-014-0157-0.

Pareto, V. (1896). Cours d'économie politique. F. Rouge.

Podpečan, Vid. 2023. "Can You Dance? A Study of Child–Robot Interaction and Emotional Response Using the NAO Robot" *Multimodal Technologies and Interaction* 7, no. 9: 85. <https://doi.org/10.3390/mti7090085>

Padhan, S., Mohapatra, A., Ramasamy, S. K., & Agrawal, S. (2023). Artificial Intelligence (AI) and Robotics in Elderly Healthcare: Enabling Independence and Quality of Life. *Cureus*, *15*(8), e42905. https://doi.org/10.7759/cureus.42905

Robles, T. F., Slatcher, R. B., Trombello, J. M., & McGinn, M. M. (2014). Marital quality and health: A meta-analytic review. *Psychological Bulletin, 140*(1), 140–187. [https://doi.org/10.1037/a0031859](https://psycnet.apa.org/doi/10.1037/a0031859)

Roelfs David J., Eran Shor, Rachel Kalish, Tamar Yogev. (2011). The Rising Relative Risk of Mortality for Singles: Meta-Analysis and Meta-Regression, *American Journal of Epidemiology*, Volume 174, Issue 4, Pages 379–389, <https://doi.org/10.1093/aje/kwr111>

Rodríguez-Martínez, A., Amezcua-Aguilar, T., Cortés-Moreno, J., & Jiménez-Delgado, J. J. (2023). Qualitative Analysis of Conversational Chatbots to Alleviate Loneliness in Older Adults as a Strategy for Emotional Health. *Healthcare (Basel, Switzerland)*, *12*(1), 62. <https://doi.org/10.3390/healthcare12010062>

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

Smith, Michael (1994). The moral problem. Cambridge, Mass., USA: Blackwell. Princeton, NJ: Princeton University Press

Sutton, R. S. & Barto, A. G. (2018). Reinforcement learning: An introduction. *Second Edition*. Cambridge, Mass: MIT Press.

Symons, J. (2021). Meaningfulness and Kinds of Normative Reasons. *Philosophia* **49**, 459–471. https://doi.org/10.1007/s11406-020-00258-7

Symons, J., & Abumusab, S. (2024). Social Agency for Artifacts: Chatbots and the Ethics of Artificial Intelligence. *Digital Society*, *3*(1)

Symons, J., & Alvarado, R. (2022). Epistemic injustice and data science technologies. *Synthese*, *200*(2),

Troshani I., Rao Hill S., Sherman C., Arthur D. (2021). Do we trust in AI? Role of anthropomorphism and intelligence. J. Comput. Inf. Syst. 61, 481–491. doi: 10.1080/08874417.2020.1788473

UC Berkeley (2020). “What is Machine learning?”. Accessed August 22, 2024. <https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/>

Umberson D, Crosnoe R, Reczek C. (2010). Social relationships and health behavior across the life course. Annual Review Of Sociology. 2010:36139–159. doi: 10.1146/annurev-soc-070308-120011.

Veronese, Nicola, Daiana Galvano, Francesca D’Antiga, Chiara Vecchiato, Eva Furegon, Raffaella Allocco, Lee Smith, Giovanni Gelmini, Pietro Gareri, and Marco Solmi (2021), "Interventions for Reducing Loneliness: An Umbrella Review of Intervention Studies," Health & Social Care in the Community, 29 (5), e89-e96

1. Bianchi and Vohs explain that income predicts the nature of social contact. “People with higher incomes spent less time with their families and neighbors and spent more time with their friends. These findings suggest that income is associated with how and with whom people spend their time” (Bianchi and Vohs 2016). Other factors that cause loneliness such as immigration, increased engagement with technology, race, educational level and so on have all been widely discussed. Given this large body of empirical evidence and variety of competing explanations, it is safe to assume that people are lonelier than they used to be decades ago in the United States. [↑](#footnote-ref-1)
2. This definition is meant to capture the core features of usage within the psychological literature. Mcgraw (1995) for example described it “as a disagreeable and debilitating way of feeling apart from others rather than being a part of them”. “It can also be described as a deficiency of the needs and metaneeds of intimacy/meaning, specifically of that kind of intimacy which is meaningful and of that meaning which is intimate” (Mcgraw, 1995). [↑](#footnote-ref-2)
3. Loneliness has been linked to several health conditions and reduced mortality rates (Cacioppo J. et. al, 2002; Robles et al., 2014, Shor & Roelfs, 2015). Mcgraw argued that human beings crave intimacy the most, which are usually entailed in love, friendship and community and to be without this is to be lonely (Mcgraw, 1995). He posits that “loneliness exhibits itself as an emotional hunger for intimacy/meaning, and extreme loneliness manifests itself as intimacy/meaning starvation” (Mcgraw, 1995). (Ma R., et al, 2020) also described loneliness as “the unpleasant experience that occurs when there is a subjective discrepancy between desired and perceived availability and quality of social interactions''. [↑](#footnote-ref-3)
4. We would like to thank an anonymous referee for suggesting the inclusion of this statement of the problem. [↑](#footnote-ref-4)
5. At first glance, both the AI and the human deceiver seem to operate under similar pretenses - they engage in relationships that are, at their core, inauthentic. However, as in Harry Frankfurt’s account of the distinction between lies and bullshit (Frankfurt, 2005), the AI, like the bullshitter is not even speaking falsely. Insofar as the bullshitter on Frankfurt’s reading is speaking in a way that is strategically designed to make his audience hold him in a high opinion without concern for truth or falsity, one might say that the AI, who has no such concern, is not even a bullshitter! [↑](#footnote-ref-5)
6. However, our biological parents are uniquely related to our identity. A human person depends in a metaphysically necessary way on their biological parents of origin. It makes no sense to say that I could have had biological parents other than those I have. [↑](#footnote-ref-6)
7. Elsewhere we have argued that an agent’s decision to favor one kind of value over others reflects that agent’s understanding of what counts as meaningful (Symons 2020). [↑](#footnote-ref-7)