The Relational Turn
Thinking Robots Otherwise

David J. Gunkel

1 Introduction

Ethics is an exclusive undertaking. In confronting and dealing with others—whether another human person, a non-human animal, or an artifact—we inevitably make a decision between *who* is worthy of consideration and respect and *what* remains a mere thing that can be used as we see fit. These decisions are often accomplished and justified on the basis of some fundamental and intrinsic property that is determined to belong to the entity by its very nature. "The standard approach to the justification of moral status is," as Mark Coeckelbergh (2012: 13) explains, "to refer to one or more (intrinsic) properties of the entity in question, such as consciousness or the ability to suffer. If the entity has this property, this then warrants giving the entity a certain moral status."

This way of proceeding has been successfully utilized on both sides of the debate concerning the moral status of AI, robots, and other artifacts. On the one side, those opposing any form of moral status for artifacts assert that these technologies are just things or objects that do not possess and will not come to possess the necessary conditions or capabilities to be considered something more. On the other side, there are those who favor extending some aspect of moral status to AI and robots by arguing that these technological things either have or will soon be able to possess one or more of the necessary and essential properties to be something other than a mere thing. What is interesting about this debate is not what makes the one side different from the other; what is interesting is what both sides already agree upon and share in order to come into conflict in the first place. And the real problem is not that this shared moral scaffolding has somehow failed to work in the face or the faceplate of AI and robots. The problem is that it has and continues to

work all-too-well, exerting its influence and operations almost invisibly and without question.

This chapter is designed to respond to this problem. It begins by first identifying and critically examining three seemingly intractable philosophical difficulties with the standard method for deciding questions of moral status. In response to these demonstrated difficulties, the second section will introduce and describe an alternative model, one which shifts the emphasis from internal properties of the individual entity to extrinsic social circumstances and relationships. The final section will then consider three possible objections to this "relational turn" and provide responses to these criticisms. The goal in all of this is not to complicate things but to introduce and formulate a metaethical theory that is more agile in its response to the unique opportunities and challenges of the 21st century.

2 The Properties Approach

In responding to others (and doing so responsibly), we typically need to distinguish between *what* is a thing and *who* is another person. As Immanuel Kant (2012: 40) once described it: "Beings whose existence rests not indeed on our will but on nature, if they are non-rational beings, still have only a relative worth, as means, and are therefore called *things*, whereas rational beings are called *persons*, because their nature already marks them out as ends in themselves, i.e., as something that may not be used merely as means, and hence to that extent limits all choice (and is an object of respect)." This just sounds intuitively correct. We go out into the world and deal with others, knowing there's a difference between other persons who are subject to respect as ends in themselves and those things that are mere objects with instrumental value as a means to an end. As Robert Esposito (2015: 1), who arguably wrote the book on this matter, explains: "If there is one assumption that seems to have organized human experience from its very beginnings it is that of a division between persons and things. No other principle is so deeply rooted in our perception and in our moral conscience…"

What is important is not this difference, but how this differentiation comes to be decided and justified. In order for something to have anything like moral or legal status, it would need to be recognized as another subject and not just an object, for example, a tool or an instrumental means. Standard approaches to addressing and resolving these matters typically

proceed, as Coeckelbergh (2012: 14) points out, by following a rather simple and straight-forward decision-making process or what could be called a moral status algorithm:

1) Having property P is sufficient for moral status S
2) Entity E has property P
3) Entity E has moral status S

In this transaction, *we* (and who is included in this first-person plural pronoun is not without consequences) first make a determination as to what ontological property or set of properties are sufficient for something to have a particular claim to moral recognition and respect. In effect, there needs to be a prior identification of what are determined to be the essential qualifying criteria that are needed for "something" to be recognized as "someone" (Spaeman 2006). We then investigate whether an entity, i.e. a robot or AI device, either currently existing or theoretically possible, actually possesses that property or set of properties (or not). Finally, and by applying the criteria decided in step one to the entity identified in step two, it is possible to "objectively" determine whether the entity in question either can or cannot have a claim to moral status or is to be regarded as a mere thing.

This way of proceeding sounds intuitively correct and natural. On this account, questions regarding moral status—decisive questions that decide where to draw the line separating things from persons—are firmly anchored in and justified by the essential nature or being of the entity that is determined to possess them. In this transaction, what something *is* determines how it *ought* to be treated. Or to put it in more formalistic terminology: ontology precedes and determines social, moral, and even legal status. But there are three problems with the approach. The first two—determination and definition—concern complications with the major premise; the third—detection—concerns problems affecting the minor premise.

2.1 Determination

How does one determine which exact property or set of properties are necessary and sufficient for something to have moral status or, as Hannah Arendt (1968: 296) puts it, "the right to have rights?" In other words, which one, or ones, count? The history of moral philosophy can, in fact, be read as something of an on-going debate and struggle over this matter with different prop-

erties vying for attention at different times. And in this process many properties—that at one time seemed both necessary and sufficient—have turned out to be either spurious, prejudicial, or both.

Take for example a rather brutal action recalled by the naturalist Aldo Leopold (1966: 237) at the beginning of his seminal essay on environmental ethics: "When god-like Odysseus, returned from the wars in Troy, he hanged all on one rope a dozen slave-girls of his household whom he suspected of misbehavior during his absence. This hanging involved no question of propriety. The girls were property. The disposal of property was then, as now, a matter of expediency, not of right and wrong." At the time Odysseus is reported to have done this, only male heads of the household were considered legitimate moral and legal subjects. Everything else—his women, his children, his animals—were property that could be disposed of without any ethical worries or reflection whatsoever. But from where we stand now, the property "male head of the household" is clearly a spurious and prejudicial criterion for determining moral status.

Similar problems are encountered with, for example, the property of rationality, which is the criterion that eventually replaces the seemingly spurious "male head of the household." When Kant (1985: 17) had defined morality as involving the rational determination of the will, non-human animals, which do not (at least since Descartes had decided that animals were mindless mechanisms) possess reason, are immediately and categorically excluded from consideration. It is because the human being possesses reason, that he—and "human being," in this case, was principally defined and characterized as male, which was the "oversight" Mary Wollstonecraft sought to address by way of her *Vindication of the Rights of Women*—is raised above the instinctual behavior of a mere brutes and able to act according to the principles of pure practical reason (Kant 1985: 63).

The property of reason, however, is contested by efforts in animal rights philosophy, which begins, according to Peter Singer, with a critical response issued by Jeremy Bentham (2005: 283): "The question is not, 'Can they reason?' nor, 'Can they talk?' but 'Can they suffer?'" For Singer, the morally relevant property is not speech nor reason, which he believes sets the bar for moral inclusion too high, but sentience and the capability to suffer. In the book *Animal Liberation* (1975) and subsequent writings, Singer argues that any sentient entity, and thus any being that can suffer, has an interest in not suffering and therefore deserves to have that interest taken into account. Tom Regan, however, disputes this determination, and focuses his "animal rights" thinking

on an entirely different property. According to Regan, the morally significant property is not rationality or sentience but what he calls "subject-of-a-life" (Regan 1983: 243). Following this determination, Regan argues that many animals, but not all animals (and this qualification is important, because the vast majority of animal are actually excluded from his brand of "animal rights"), are "subjects-of-a-life": they have wants, preferences, beliefs, feelings, etc. and their welfare matters to them (Regan 1983). Although these two formulations of animal rights effectively challenge the anthropocentric tradition in moral philosophy, there remains disagreements about which exact property is the necessary and sufficient condition for moral consideration.

2.2 Definition

Irrespective of which property (or set of properties) is selected, they each have problems with definition. Take, for example, the property of consciousness, which is often utilized in the discussions and debates regarding moral status for intelligent machines and artifacts. Unfortunately, there is no univocal and widely accepted definition. The problem, as Max Velmans (2000: 5) points out, is that the term unfortunately "means many different things to many different people, and no universally agreed core meaning exists." In fact, if there is any general agreement among philosophers, psychologists, cognitive scientists, neurobiologists, AI researchers, and robotics engineers regarding the property of consciousness, it is that there is little or no agreement when it comes to defining and characterizing the concept. Although consciousness, as the theologian Anne Foerst remarks, is the secular and supposedly more "scientific" replacement for the occultish "soul", it turns out to be just as much an occult property (Benford/Malartre 2007: 162).

Other essential properties do not do much better. Suffering and the experience of pain—which is the property usually deployed in non-standard patient-oriented approaches like animal rights philosophy—is just as problematic, as Daniel Dennett demonstrates in "Why You Cannot Make a Computer that Feels Pain." In this provocatively titled essay, Dennett imagines trying to disprove the standard argument for human (and animal) exceptionalism "by actually writing a pain program, or designing a pain-feeling robot" (Dennett 1998: 191). At the end of what turns out to be a rather protracted and detailed consideration of the problem, Dennett concludes that we cannot, in fact, make a computer that feels pain. But the reason for drawing this conclusion does not derive from what one might expect. According to Dennett, the

reason you cannot make a computer that feels pain is not the result of some technological limitation with the mechanism or its programming. It is due to the fact that we remain unable to decide what pain is in the first place. What Dennett demonstrates, therefore, is not that some workable concept of pain cannot come to be instantiated in the mechanism of a computer or a robot, either now or in the foreseeable future, but that the very concept of pain that would be instantiated is already arbitrary, inconclusive, and lacking a clear definition.

2.3 Detection

Most (if not all) of the properties that are considered morally relevant, like the experience of pain or other emotions, are internal mental states or capabilities that are not immediately accessible or directly observable. As Janina Loh (2021: 109) points out, this epistemic uncertainty is not something that is limited to AI and robots: "Actually, we do not only not know what it is like to be a machine or an animal, e.g. a bat—to quote the title of a famous paper by Thomas Nagel (1974). But we also don't really know what it is like to be another human. Because it cannot be determined with unambiguity whether humans are actually equipped with freedom of will and similar abilities. We cannot clearly prove them empirically."

This epistemological barrier is what philosophers commonly call "the problem of other minds." But it is not just a problem for philosophers, as the Brazilian anthropologist Eduardo Viveiros de Castro (2017: 52) explains: "The theological problem of the soul of others became the philosophical puzzle of 'the problem of other minds,' which currently extends so far as to include neurotechnological inquiries on human consciousness, the minds of animals, the intelligence of machines." Although this problem is not necessarily intractable, as Steve Torrance (2014) and others have argued, the fact of the matter is we cannot, as Donna Haraway (2008: 226) characterizes it, "climb into the heads of others to get the full story from the inside." Even advanced neuroimaging technology like functional magnetic resonance imaging (fMRI) does not provide an easy resolution to this epistemological uncertainty. "This type of technology," as Fabio Tollon (2021: 153) explains, "allows us to peer into the 'moving parts' in the brain which may be correlated with sentience. However, talk of internal states and the talk of how we describe, scientifically, the information that an fMRI machine represents to us are two very different language games."

Responses to this problem, then, typically rely on and mobilize behavioral demonstrations like that devised by Alan Turing for his game of imitation, which inferred machine cognition (an internal state-of-mind) from a demonstration of convincing conversational behavior (an external performance). Even if the behaviors are reasonably convincing, it is still a matter of inferring an internal cause from apparent external behaviors. "We are," as John Basl and Joseph Bowen (2020: 298) explain, "in an epistemologically poor place when it comes to determining what the preferences of an AI are, or what makes it suffer, what it may enjoy, and so on, even if we imagine that the AI is telling us what it 'likes, enjoys, desires, etc.' and behaves accordingly. This is because whatever evidence these behaviors generate is screened off by the fact that the AI might be programmed to behave that way. Yes, the AI convincingly emotes, but it also might have been designed specifically to trick us into thinking it has mental states and emotes because of that despite having no such states."

Consequently, "there is," as Dennett (1998: 172) concludes, "no proving that something that seems to have an inner life does in fact have one." Although philosophers, psychologists, cognitive scientists, and neuroscientists throw an impressive amount of argumentative and experimental effort at the problem, so far it has not been resolvable in any way approaching what would pass for definitive evidence, strictly speaking. In other words, no matter what property is identified, it is always possible to seed reasonable doubt concerning its actual presence. If an AI or a robot, for example, appears to be conscious and therefore a subject of moral concern, all that is necessary to disarm this inference is to point out that it is at least possible that what appears to be conscious behavior is in fact just an effect of clever programming. Likewise, if one seeks to exclude AI or robots from moral consideration on the grounds that they are just things that do not possess real cognitive capabilities, all that is necessary to counter this assertion is to point out that this statement might be true for current systems but may not hold for future systems that have the potential for (and even a high probability of) achieving these threshold conditions. Even if the problem of other minds is not the intractable philosophical dilemma that is often advertised, it is sufficient for sowing doubt about the presence or absence of the qualifying criteria and, by extension, rending decisions about moral status tentative, indeterminate, and uncertain.

3 The Relational Turn

In response to these problems, philosophers—especially in the continental and feminist STS traditions—have advanced other methods for resolving the question of moral status that can be characterized as a relational turn in ethics. This alternative has (at least) three pivotal characteristics:

3.1 Relational

Moral status is decided and conferred not on the basis of subjective or internal properties determined in advance but according to objectively observable, extrinsic social relationships. "Moral consideration," as Coeckelbergh (2010: 214) describes it, "is no longer seen as being 'intrinsic' to the entity: instead it is seen as something that is 'extrinsic': it is attributed to entities within social relations and within a social context." As we encounter and interact with others—whether they be another human person, a non-human animal, or a seemingly intelligent machine—it is first and foremost experienced in relation to us. Consequently, the question of moral status does not depend on what the other is in its essence but on how she/he/it (and pronouns matter here) stands in relationship to us and how we decide, in the face of the other (to use Levinasian terminology), to respond. In this transaction "relations are prior to the things related" (Callicott 1989: 110) such that, as Karen Barad (2007: 136-7) has argued, the relationship comes first—in both temporal sequence and status—and takes precedence over the individual relata.

This shift in perspective—a shift that inverts the standard operating procedure by putting the ethical relationship before determinations of ontological conditions—is not just a theoretical proposal; it has, in fact, been experimentally confirmed in numerous social science investigations. The computer as social actor (CASA) studies undertaken by Byron Reeves and Clifford Nass (1997), for example, demonstrated that human users will accord computers and other technological artifacts social standing similar to that of another human person and that this occurs as a product of the extrinsic social interaction, irrespective of the intrinsic properties (actually known or not) of the individual entities involved. Social standing, in other words, is a mindless operation. And these results have been verified in "robot abuse studies," where HRI (human robot interaction) researchers have found that human subjects respond emotionally to robots and express empathic concern for the machines irrespective of the cognitive properties or inner workings of the device.

3.2 Phenomenological

This alternative is phenomenological or (if you prefer) radically empirical in its epistemological commitments. Because moral status is dependent upon extrinsic social circumstances and not internal properties[1], the seemingly irreducible problem of other minds is not some fundamental epistemological limitation that must be addressed and resolved prior to decision making. Instead of being derailed by the epistemological problems and complications of other minds, the relational turn immediately affirms and acknowledges this difficulty as the basic condition of possibility for ethics as such. Consequently, "the ethical relationship," as Emmanuel Levinas (1987: 56) writes, "is not grafted on to an antecedent relationship of cognition; it is a foundation and not a superstructure…It is then more cognitive than cognition itself, and all objectivity must participate in it." It is for this reason that Levinasian philosophy focuses attention not on other minds, but on the *face* of the other. Or as Richard Cohen (2001: 336) succinctly explains in what could be an advertising slogan for Levinasian thought: "Not other 'minds,' mind you, but the 'face' of the other, and the faces of all others."[2]

This also means that the order of precedence in moral decision making should be reversed. Internal properties do not come first and then moral respect follows from this ontological fact. We have things backwards. We project the morally relevant properties onto or into those others who we have already decided to treat as being socially and morally significant. In social situations, then, we always and already decide between *who* counts as morally significant and *what* does not and then retroactively justify these actions by "finding" the essential properties that we believe motivated this decision-making in the first place. Properties, therefore, are not the intrinsic *a priori* condition of possibility for moral status. They are *a posteriori* products of extrinsic social interactions with and in the face of others.

---

1 The only property that would be necessary for something (like a rock) to be in relation to something else (like me) is the minimal ontological condition of being.

2 "Face" in Levinas is not a substantive property that is possessed by an entity. It is (or takes place as) an act or event of "facing." For more on this and its significance for interpretations and applications of Levinasian philosophy, see Silvia Benso's *The Face of Things* (2000).

3.3 Diverse

Finally, making moral status dependent on consciousness or other cognitive capabilities belonging to the individual is thoroughly Cartesian. Other cultures, distributed across time and space, do not divide-up and make sense of the diversity of being in this arguably binary fashion. They perform decisive cuts separating the *who* from the *what* according to other ways of seeing, valuing, and acting. Following the insights of Josh Gellers (2020), we can identify alternative ways of organizing social relationships by considering cosmologies that are not part of the Western philosophical lineage. As Archer Pechawis explains in his contribution to the essay "Making Kin with Machines":

> *nēhiyawēwin* (the Plains Cree language) divides everything into two primary categories: animate and inanimate. One is not 'better' than the other, they are merely different states of being. These categories are flexible: certain toys are inanimate until a child is playing with them, during which time they are animate. A record player is considered animate while a record, radio, or television set is inanimate. But animate or inanimate, all things have a place in our circle of kinship or *wahkohtowin* (Lewis et al. 2018).

This alternative formulation runs counter to the dominant ways of thinking, seeing the boundary between what Western ontologies call "person" and "thing" as being endlessly flexible, permeable, and more of a continuum than an exclusive opposition.

Similar opportunities/challenges are available by way of other non-Western religious and philosophical traditions. In her investigation of the social position of robots in Japan, Jennifer Robertson (2014: 576) finds a remarkably different way of organizing the difference between living persons and artificially designed/manufactured things:

> "*Inochi*, the Japanese word for 'life,' encompasses three basic, seemingly contradictory but inter-articulated meanings: a power that infuses sentient beings from generation to generation; a period between birth and death; and, most relevant to robots, the most essential quality of something, whether organic (natural) or manufactured. Thus robots are experienced as 'living' things. The important point to remember here is that there is no ontological pressure to make distinctions between organic/ inorganic, animate/inanimate, human/nonhuman forms. On the contrary, all of these forms are linked to form a continuous network of beings."

These are not the only available alternatives, and, by citing these two instances, the intention is not to suggest that these different ways of thinking difference differently are somehow "better" than those developed in Western philosophical and religious traditions. In fact (and this is where things get really complicated), making and operating on that kind of assumption would itself be an instance of "orientalism" (Said 1994), which always sought new resources derived from exoticized Others in order to rehabilitate and ensure the continued success of Western hegemony. The alternatives, by contrast, are just different and, in being different, offer the opportunity for critically questioning what is assumed to be true and often goes by without saying. Gesturing in the direction of other ways of thinking and being can have the effect of shaking one's often unquestioned confidence in cultural constructs that are already not natural, universal, nor eternally true.

4 Objections and Replies

The relational turn introduces an alternative that supplies other ways of responding to and taking responsibility for others and other forms of otherness. But it is by no means a panacea or some kind of moral theory of everything. It just arranges for other kinds of questions and modes of inquiry that are seemingly more attentive to the exigencies of life as it is encountered here and now at the beginning of the 21st century. Having said that, it is important to recognize that relational ethics is not without its own set of unique challenges—three in particular.

4.1 Relativism

For all its opportunities, the relational turn risks exposure to the charge of moral relativism, or as Charles Ess (1996: 204) explains "the claim that no universally valid beliefs or values exist." To put it rather bluntly, if moral status is "relational" and open to different decisions concerning others made at different times for different reasons, are we not at risk of affirming an extreme form of moral relativism? Versions of this objection have been brought by a number of critics, including Vincent Müller (2021) and Kęstutis Mosakas (2021). In fact, Mosakas has provided rather extensive diagnosis of the perceived problem in his contribution to John-Stewart Gordon's *Smart Technologies and Fundamental Rights*:

> As Simon Kirchin explains, "the key relativistic thought is that the something that acts as a standard will be different for different people, and that all such standards are equally authoritative" (Kirchin 2012: 15). Particularly problematic is the extreme version, which denies there being any moral judgments or standards that could be objectively true or false (in contrast to moderate versions that do admit of a certain degree of objectivity) (Moser and Carson 2001: 3). Given the apparent rejection of any such standard by Coeckelbergh and Gunkel, they seem to be hard-pressed to explain how the radically relational ethics (to use Coeckelbergh's own term (Coeckelbergh 2010: 218)) that they are advocating avoids the extreme version (Mosakas 2021: 95).

The perceived problem with relativism (especially the extreme version of it) is that it encourages and supports a situation where anything goes and all things are permitted. But as both Coeckelbergh and I have argued in other contexts (Gunkel 2018 and Coeckelbergh 2020), this particular understanding of "relative" is limited and the product of a culturally specific understanding of and expectation for ethics.

Robert Scott (1967), for instance, understands "relativism" entirely otherwise—as a positive rather than negative term: "Relativism, supposedly, means a standardless society, or at least a maze of differing standards, and thus a cacophony of disparate, and likely selfish, interests. Rather than a standardless society, which is the same as saying no society at all, relativism indicates circumstances in which standards have to be established cooperatively and renewed repeatedly" (Scott 1967: 264). Chares Ess (2009: 21) calls this alternative "ethical pluralism." "Pluralism stands as a third possibility—one that is something of a middle ground between absolutism and relativism… Ethical pluralism requires us to think in a 'both/and' sort of way, as it conjoins both shared norms and their diverse interpretations and applications in different cultures, times, and places." Likewise, Luciano Floridi (2013: 32) advocates a "pluralism without endorsing relativism," calling this third alternative or "middle ground" relationalism.

Others, like Rosi Braidotti, call upon and mobilize "a form of non-Western perspectivism," which exceeds the grasp of Western epistemology. "Perspectivism," as Viveiro de Castro (2015: 24) explains in his work with Amerindian traditions, "is not relativism, that is the affirmation of the relativity of truth, but relationalism, through which one can affirm *the truth of the relative is the relation*.'" For Braidotti (2019: 90) perspectivism is not just different from but is "the antidote to relativism." "This methodology," as she explains, "respects dif-

ferent viewpoints from equally materially embedded and embodied locations that express the degree of power and quality of experience of different subjects." Braidotti therefore recognizes that what is called "truth" is always formulated and operationalized from a particular subject position, which is dynamic, different, and diverse. The task is not to escape from these differences in order to occupy some fantastic transcendental vantage point but to learn how to take responsibility for these inescapable alterations in perspective and their diverse social, moral, and material consequences. The relational turn, therefore, does not endorse relativism (as it is typically defined) but embodies and operationalizes an ethical pluralism, relationalism, or perspectivism that complicates the simple binary logic that defines relativism in opposition to moral absolutism.

4.2 Dehumanization

If moral status is not substantiated by ontological properties but is the product of external relations with others, doesn't this run the risk, as noted by Anne Gerdes (2015: 274), that we might lose something valuable, that "our human-human relations may be obscured by human-robot relations?" This is precisely the concern of Kathleen Richardson (2019: 1), who argues that the relational turn is just as likely to be (mis)used to instrumentalize and reify human subjects: "But if the machine can become another, what does it say for how robotic and AI scientists conceptualise 'relationship'? Is relationship instrumental? Is relationship mutual and reciprocal?" What worries Richardson is that relational ethics—with its focus on social embedding and externalities—would allow for and justify treating other humans as things and not persons. And her response to and fix for this potential hazard is to retreat to a dogmatic reassertion of human exceptionalism. "Humans are never tools or instruments, even if relations between people take on a formal character…In every encounter we meet each other as persons, members of a common humanity" (Richardson 2019: 1).

A different way to respond to this challenge is to recognize, as Anne Foerst suggests, that otherness is not (not in actual practices, at least) unlimited or absolute: "Each of us only assigns personhood to a very few people. The ethical stance is always that we have to assign personhood to everyone, but in reality we don't. We don't care about a million people dying in China of an earthquake, ultimately, in an emotional way. We try to, but we can't really, because we don't share the same physical space. It might be much more im-

portant for us if our dog is sick" (Benford/Malartre 2007: 163). This statement is perhaps more honest about the way that moral decision-making and the occurrence of otherness actually transpires. Instead of declaring an absolutist claim to a kind of dogmatic totality, it remains open to particular configurations of otherness that is mobile, flexible, and context-dependent. It is a kind of posthumanist ethic that, as Barad (2007: 136) describes it "doesn't presume the separateness of any-'thing,' let alone the alleged spatial, ontological, and epistemological distinctions that set humans apart."

But there remains, as Gerdes (2015) insightfully recognizes, something in this formulation that is seemingly abrasive to our moral intuitions. This may be due to the fact that this way of thinking does not make a singular and absolute decision about otherness that stands once and for all, such that there is one determination concerning others that decides everything for all time. The encounter with others—the occurrence of face in the face of the other—is something that happens in time and needs to be negotiated and renegotiated. This means that the work of ethics is ostensibly inexhaustible. It is an ongoing and interminable responsibility requiring that one respond and take responsibility for how one responds. Is this way of thinking and doing ethics without risk? Not at all. But the risk is itself the site of ethics and the challenge that one must face in all interactions with others, whether human, animal, or otherwise (cf. Gunkel 2012).

4.3 Performative Contradiction

For all its rhetorical posturing, the relational turn still seems to be inescapably anthropocentric and dependent on properties. This objection is something that is developed by Henrik Skaug Sætra in the essay "Challenging the Neo-Anthropocentric Relational Approach to Robot Rights." Sætra directs his critical efforts to what he identifies as two performative contradictions, where what is espoused by the relational turn betrays or appears to be inconsistent with what it actually does. "Relationalism," as he (2021: 1) explains, "purportedly opens the door for considering robot rights and moving past anthropocentrism. However, I argue that relationalism is, quite to the contrary, a form of neo-anthropocentrism that re-centers human beings and their unique ontological properties, perceptions, and values."

The first critical target in this effort is anthropocentrism or (better stated) its opposite. According to Sætra's reading of the literature, the relational turn promotes itself as being non-anthropocentric but actually is anthropocentric

in practice. "My first objection is that relationalism is arguably deeply anthropocentric because moral standing is derived exclusively from how human beings perceive and form relations with other entities. As we have seen, moral standing is here derived from how something is treated, and not what it is. This means that humans are key to determining value, as it is how entities are treated and perceived by humans that determine their moral standing" (Sætra 2021: 6). This criticism is correct. Or, better stated, it is not at all a criticism but an accurate description and characterization.

First, "humans are key to determining value" insofar as morality is a human endeavor. Faulting relationalism for being organized and directed by concerns about human interests and values would be like faulting physics or chemistry for not developing a purely objective science that is not at all involved with human perceptions, concepts, and modes of understanding. The relational turn, like all forms of what Donna Haraway (1991) calls "situated knowledge," comes from somewhere and is embedded and embodied in a specific subject position. To expect that any form of human knowledge would be able to escape from these human-all-too-human conditions of possibility and operate from some super-human position of transcendental objectivity is a metaphysical fantasy that is reserved for the gods. In other words, the axiological purity that Sætra operationalizes as a kind of "litmus test" is a metaphysical fantasy. So yes, the relational turn, like all moral theories and practices, is dependent upon human capabilities, perceptions, and values. And, like all sciences, the critical task is not and cannot be to escape from these existential preconditions but to learn how to respond to and to take responsibility for them.

Second, Sætra is right to conclude from this that relationalism is *not* non-anthropocentric. But he is incorrect in concluding that this double negative implies a positive, namely that it is morally anthropocentric. Non-anthropocentric ethical theories, as Sætra characterizes and explains, include a number of moral innovations that aim to decenter human exceptionalism and culminate, for him at least, in ethical biocentrism: "As compared to the previous type of non-anthropocentrism, ethical biocentrism does not require us to uncover, or conjure up, the interests, preferences, etc., of other entities. Instead, they are considered valuable just because of being what they are, which is why the terms intrinsic or inherent value are often used" (Sætra 2021: 5). But this is not—at least in terms of logical structure—really all that different. Like the anthropocentric model that it contests, ethical biocentrism is still an ontology-driven transaction, where what something is—its being what it is—de-

termines intrinsic or inherent value. The problem, then, is not just with anthropocentrism or its alternatives, but with any and all "epistemic centrisms" which, as Loh (2021: 109) points out "remain committed to the paternalism implicit in the subject-object dichotomy."

The relational turn does not play by these rules. It deliberately flips the script on this entire metaphysical transaction. Following the innovations of Levinas (1969: 304), who famously overturned 2000+ years of Western philosophy by proclaiming that ethics is first philosophy, the relational turn puts the moral relationship *first* in terms of both sequence and status. Or as Barad (2007: 139) describes it, the primary unit "is not independent objects with inherent boundaries and properties," but relations—"relations without preexisting relata."[3] This fundamental change in perspective produces something outside the orbit of either anthropocentrism or its non-anthropocentric others, producing an "eccentric moral theory" (Gunkel 2018: 164) that deconstructs the very difference that distinguishes anthropocentrism from its various alternatives.

This leads to the second critical target, which concerns the status and function of properties. "My second objection," Sætra (2021: 7) writes, "is that relationalism is in reality a camouflaged variety of the properties-based approach. This is so because how we relate to other entities is determined by the properties of these others." In other words, the relational turn can say that it puts relations before relata and makes determinations about moral status dependent on "how something is treated, and not what it is" (Sætra 2021: 6). But this is just patently false, because properties still matter. "How we relate to someone, and how an entity acts, is dependent on their properties." Again Sætra right, but not for the right reasons.

Properties do play a role in moral decision making, and they can be a useful and expedient heuristic for responding to and taking responsibility in the face of others. What is at issue, then, is not their importance but their status and function. As Gellers points out, properties are not antithetical to or excluded from relationalism, they are just recontextualized and understood

3 This does not mean that all things are essentially nothing outside of being relata. The thing-in-itself (to use Kantian terminology) is ontologically consistent in and of itself. The thing as it stands in relationship to another—as relata—is dependent upon the terms and conditions of the relationship. That fact does not mean (continuing with the Kantian formulation) that there is no *Ding an sich*. It is not nothing; it is just epistemologically inaccessible as it is in itself.

in relational terms. "Coeckelbergh," Gellers (2020: 19) writes, "does not foreclose the possibility that properties may play a role in a relational approach to moral consideration. Instead, he leaves room for "properties-as-they-appear-to-us within a social-relational, social-ecological context (Coeckelbergh 2010: 219)." In other words, the properties that are determined to belong to an entity are actually a phenomenal effect of the relationship and not an antecedent ontological condition and cause. This flips the script on things.

In moral philosophy—at least its standard Western varieties—what something is commonly determines how it ought to be treated. Or as Luciano Floridi (2013: 116) describes it: "what the entity is determines the degree of moral value it enjoys, if any." According to this largely unchallenged standard operating procedure, the question concerning the status of others—whether they are someone who matters or something that does not—is entirely dependent on and derived from what it is and what capabilities it possesses. Ontology, therefore, is *first* in both procedural sequence and status. Sætra not only endorses this way of thinking but normalizes and naturalizes it, even though it is the product of a specific philosophical tradition and culture.

The relational turn not only challenges this way of thinking but deliberately reverses its procedure. This does not diminish the role of properties, it simply inverts the direction of the derivation. The morally significant properties—those ontological criteria that we had assumed grounded moral respect—are actually what Slavoj Žižek (2008: 209) calls "retroactively (presup)posited" as the result of and as justification for prior decisions made in the face of social interactions with others. Consequently, even before we know anything at all about what something is in its essence, we have already been called upon and obligated to make a decisive response.[4]

To give it a Kantian spin, we can say that what something is in itself—*das Ding an sich*—is forever inaccessible insofar as all we ever have access to is how something appears to be relative to us. Whatever we think it is in-itself is the result of something we project onto or into it after the fact. So it is not accurate to conclude that "relationalism is in reality a camouflaged variety of the properties-based approach." Such a conclusion is possible if and only if one normalizes and naturalizes the standard derivation of "ought" from "is." It is just as likely—and maybe even more epistemologically honest—to conclude

4 For this reason, relations are neither an ontological criterion nor an epistemic category. They are the prior ethical condition. This is why, for Levinas (1969, 304), it is ethics, and not ontology or epistemology, that is "first philosophy."

that what is actually an effect of embedded and embodied interactions with others has been mistakenly dressed-up and masquerading as a cause.

5 Conclusions

The question concerning the moral status of AI and robots is not really about the artifact. It is about us and the limits of who is included in and what comes to be excluded from that first-person plural pronoun, "we." It is about how *we* decide—together and across differences—to respond to and take responsibility for our shared social reality. It is, then, in responding to the moral opportunities and challenges posed by seemingly intelligent and social artifacts that we are called to take responsibility for ourselves, for our world, and for those others who are encountered here.

In devising responses to these challenges, we can obviously deploy the standard properties approach. This method has the weight of history behind it and therefore constitutes what can be called the default setting for addressing questions concerning moral status. It is, to use the terminology of Thomas Kuhn (1996), widely accepted as "normal science." But this normalized approach, for all its advantages, also has demonstrated difficulties with the determination, definition, and detection of the qualifying essential properties. This does not mean, it is important to point out, that the properties approach is somehow wrong, misguided, or refuted on this account. It just means that this way of thinking—despite its almost unquestioned acceptance as normal science within Western traditions—has limitations and that these limitations are becoming increasingly evident in the face or the faceplate of AI and robots—in the face of others who are and remain otherwise.

To put it in terms borrowed from Žižek (2006), the properties approach, although appearing to be the right place to begin thinking about and resolving the question of machine moral standing, may turn out to be the "wrong question" and even an obstacle to its solution. As an alternative, the relational turn formulates an approach to addressing the question of moral status that is situated and oriented otherwise. This alternative circumvents many of the problems encountered in the properties approach by arranging for an ethics that is relational, phenomenological, and diverse. Whether this alternative provides for a better way to formulate moral decisions is something that will need to be determined and decided in the face of others.

References

Arendt, Hannah (1968): The Origins of Totalitarianism, New York: Harcourt Books.

Barad, Karen (2007): Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning, Durham, NC: Duke University Press.

Basl, John/Bowen, Joseph (2020): "AI as a Moral Right-Holder." In: Markus D. Dubber/Frank Pasquale/Sunit Das (eds.), The Oxford Handbook of Ethics of AI, New York: Oxford University Press, pp. 289-306.

Benford, Gregory/Malartre, Elisabeth (2007): Beyond Human: Living with Robots and Cyborgs, New York: Tom Doherty.

Benso, Silva (2000): The Face of Things: A Different Side of Ethics. Albany, NY: SUNY Press.

Bentham, Jeremy (2005 [1789]): An Introduction to the Principles of Morals and Legislation. New York: Oxford University Press.

Braidotti, Rosi (2019): Posthuman Knowledge, Cambridge: Polity Press.

Callicott, J. Baird (1989): In Defense of the Land Ethic: Essays in Environmental Philosophy, Albany, NY: State University of New York Press.

Coeckelbergh, Mark (2010): "Robot Rights? Towards a Social-Relational Justification of Moral Consideration." In: Ethics and Information Technology 12/3, pp. 209-221.

Coeckelbergh, Mark (2012): Growing Moral Relations: Critique of Moral Status Ascription, New York: Palgrave Macmillan.

Cohen, Richard (2001): Ethics, Exegesis, and Philosophy: Interpretation After Levinas, Cambridge: Cambridge University Press.

Dennett, Daniel (1998): Mindstorms, Cambridge, MA: MIT Press.

Esposito, Roberto (2015): Persons and Things: From the Body's Point of View. Translated by Zakiya Hanafi, Cambridge: Polity Press.

Ess, Charles (1996): "The Political Computer: Democracy, CMC, and Habermas." In: Charles Ess (ed.), Philosophical Perspectives on Computer-Mediated Communication, Albany, NY: SUNY Press, pp. 196-230.

Ess, Charles (2009): Digital Media Ethics, Cambridge: Polity Press.

Floridi, Luciano (2013): The Ethics of Information, New York: Oxford University Press.

Gellers, Joshua C. (2020): The Rights of Robots: Artificial Intelligence, Animal and Environmental Law, New York: Routledge.

Gerdes, Anne (2015): "The Issue of Moral Consideration in Robot Ethics." In: ACM SIGCAS Computers & Society 45/3, pp. 274-280.

Gunkel, David J. (2012): The Machine Question: Critical Perspectives on AI, Robots and Ethics. Cambridge, MA: MIT Press.

Gunkel, David J. (2018): Robot Rights, Cambridge, MA: MIT Press.

Haraway, Donna (1991): Simians, Cyborgs and Women: The Reinvention of Nature, New York: Routledge.

Haraway, Donna (2008): When Species Meet, Minneapolis, MN: University of Minnesota Press.

Kant, Immanuel (1985 [1788]): Critique of Practical Reason. Translated by Lewis White Beck, New York: Macmillan.

Kant, Immanuel (2012 [1786]): Groundwork of the Metaphysics of Morals. Translated by Mary Gregor and Jens Timmermann, Cambridge: Cambridge University Press.

Kirchin, Simon (2012): Metaethics, New York: Palgrave Macmillan.

Kuhn, Thomas S. (1996 [1962]): The Structure of Scientific Revolutions, Chicago: University of Chicago Press.

Leopold, Aldo (1966): A Sand County Almanac, New York: Oxford University Press.

Levinas, Emmanuel (1969): Totality and Infinity: An Essay on Exteriority. Translated by Alphonso Lingis, Pittsburgh, PA: Duquesne University.

Levinas, Emmanuel (1987): Collected Philosophical Papers. Translated by Alphonso Lingis: Dordrecht: Martinus Nijhoff.

Lewis, Jason Edward/Arista, Noelani/Pechawis, Archer/Kite, Suzanne (2018): "Making Kin with the Machines." In: Journal of Design and Science 16; https://doi.org/10.21428/bfafd97b.

Loh, Janina (2021): "Ascribing Rights to Robots as Potential Moral Patients." In: John-Stewart Gordon (ed.), Smart Technologies and Fundamental Rights, Leiden and Boston: Brill Rodopi, pp. 101-126.

Mosakas, Kęstutis (2021): "Machine Moral Standing: In Defense of the Standard Properties-Based View." In: John-Stewart Gordon (ed.), Smart Technologies and Fundamental Rights, Leiden and Boston: Brill Rodopi, pp. 73-100.

Moser, Paul K./Carson, Thomas L. (2001): Moral Relativism: A Reader, New York: Oxford University Press.

Müller, Vincent C. (2021): "Is it Time for Robot Rights? Moral Status in Artificial Entities." In: Ethics and Information Technology 23, pp. 579-587.

Nagel, Thomas (1974): "What's it Like to be a Bat." In: The Philosophical Review 83/4, pp. 435-450.

Reeves, Byron/Nass, Clifford (1996): The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places, Cambridge: Cambridge University Press.

Regan, Tom (1983): The Case for Animal Rights, Berkeley, CA: University of California Press.

Richardson, Kathleen (2019): Rethinking the I-You Relation Through Dialogical Philosophy in the Ethics of AI and Robotics. In: AI & Society 34, pp. 1-2.

Robertson, Jennifer (2014): "Human Rights vs. Robot Rights: Forecasts from Japan." In: Critical Asian Studies 46/4, pp. 571-598.

Sætra, Henrik Skaug (2021): "Challenging the Neo-Anthropocentric Relational Approach to Robot Rights." In: Frontiers in Robotics and AI 8, pp. 1-9.

Said, Edward (1994 [1978]). Orientalism, New York: Vintage Books.

Scott, Robert L. (1967): "On Viewing Rhetoric as Epistemic." In: Central States Speech Journal 18, pp. 9-17.

Singer, Peter (1975): Animal Liberation: A New Ethics for Our Treatment of Animals, New York: New York Review of Books.

Spaemann, Robert (2006): Persons: The Difference between "Someone" and "Something." Translated by Oliver O'Donovan, New York: Oxford University Press.

Tollon, Fabio (2021): "The Artificial View: Toward a Non-Anthropocentric Account of Moral Patiency." In: Ethics and Information Technology 23/1, pp. 147-155.

Torrance, Steve (2014): "Artificial Consciousness and Artificial Ethics: Between Realism and Social Relationism." In: Philosophy & Technology 27/1, pp. 9-29.

Velmans, Max (2000): Understanding Consciousness, New York: Routledge.

Viveiros de Castro, Eduardo (2015): The Relative Native: Essays on Indigenous Conceptual Worlds. Translated by Martin Holbraad, David Rodgers and Julia Sauma, Chicago: HAU Press.

Viveiros de Castro, Eduardo (2017): Cannibal Metaphysics: For a Post-Structural Anthropology. Translated by Peter Skafish, Minneapolis, MN: University of Minnesota Press.

Wollstonecraft, Mary (2004 [1792]): A Vindication of the Rights of Woman. New York: Penguin Classics.

Žižek, Slavoj (2006): "Philosophy, the 'Unknown Knowns,' and the Public Use of Reason." In: Topoi 25, pp. 137-142.

Žižek, Slavoj (2008): For They Know Not What They Do: Enjoyment as a Political Factor, London: Verso.