

Thinking unwise: a relational u-turn

Nicholas Barrow. Accepted submission to *RoboPhilosophy2022*

Abstract:

In this paper, I add to the recent flurry of research concerning the moral patiency of artificial beings. Focusing on David Gunkel's adaptation of Levinas, I identify and argue that the Relationist's extrinsic case-by-case approach of ascribing artificial moral status fails on two accounts. Firstly, despite Gunkel's effort to avoid anthropocentrism, I argue that Relationism is, itself, anthropocentric in virtue of how its case-by-case approach is, necessarily, assessed from a human perspective. Secondly I, in light of interpreting Gunkel's Relationism as a case-by-case approach, reiterate that it cannot give sufficient action guidance.

1. Introduction

Who, or what, do we owe equal treatment? Initially restricted to “other men”, our answers now include “foreigners, women, animals, and even the environment” (Gunkel, 2012 p.1). Our next frontier, however, is product of our evolution: artificial beings. This latest challenge raises issues old and new. Much like with animals (Singer, 1981), if we are wrong in treating artificial beings indifferently, we may be causing unimaginable suffering (Andreotta, 2021). Moreover, in practice, if we ascribe an entity artificial moral status (AMS), we must act accordingly. AMS also raises *new* issues - that of responsibility. For instance, if AMS and

responsibility are linked, then what we ascribe AMS matters greatly to whom is held accountable for a robot's actions. If we wrongly ascribe AMS, a robot's producers could avoid accountability by inappropriately shifting the blame to the robot itself (Véliz, 2021). The philosophical and societal stakes of correctly discerning AMS are prevalent.

Broadly, theorists of AMS ascription fall into two camps: those that focus on properties, and those that think *otherwise*. The former ascribes AMS ontologically, requiring an entity to *have* certain intrinsic criteria. Whereas, for the latter, "it does not matter what's going on, 'on the inside'" (Danaher, 2020 p.2025). Thinking otherwise theorists can be further divided into Behaviourists (Danaher, 2021; Danaher, 2020, Coeckelbergh, 2010a), and Relationists (Gunkel, 2018a; Gunkel, 2018b; Coeckelbergh, 2010b). For the Behaviourist, an artificial entity need not *have* particular properties, it must merely *behave* as another that does.

In this paper, however, I will critically assess the latter's method of ascription: that AMS should be granted in virtue of extrinsic, social-relational context. I will argue that the Relationist's method of ascription fails on two accounts. Firstly, despite their effort to avoid anthropocentric, unfair, methods of ascription, the Relationist's *own* method is anthropocentric. This is in virtue of their case-by-case assessment of the *other* - where the assessor is *necessarily* human. Secondly, following this case-by-case account, even if the Relationist conceded anthropocentrism, they provide insufficient action guidance. Ultimately, I will conclude that not only is the Relationist method of ascribing AMS anthropocentric – that they themselves argue against – but Relationism also collapses into moral relativism and, thus, cannot give sufficient action guidance.

2. The relational turn

2.1 Against Property-Based Approaches

Standard methods of ascribing AMS can be referred to as property-based approaches. They take the form: for x to be ascribed AMS, x must *have* property x , y , z , etc. Property-based conceptions of AMS, therefore, commit to an ontological exploration of which properties are intrinsically morally relevant. Properties, on this standard view, are “intrinsic *a priori* condition(s) of possibility” for AMS (Gunkel, 2018a §6.1.3). Typical candidates include Sentience (Véliz, 2021; Eskens, 2015), Consciousness (Andreotta, 2021; Mosakas, 2020), and Autonomy (Hakli and Mäkelä, 2019; Sullins, 2006; Floridi and Sanders, 2004).

Gunkel raises multiple challenges to the standard, property-based approach of ascribing AMS to motivate his own, social-relational view (Gunkel, 2018a §3.3.2). I will briefly outline two of them:

Firstly, which property do we choose? There is an array of disagreement - I have mentioned Sentience, Consciousness, and Autonomy, however, this is not an exhaustive list. If we cannot agree on what properties matter for AMS, Gunkel argues, it follows that property-based approaches are not a good grounding for AMS. Secondly, even if we did agree on *which* properties are relevant, there remains disagreement on what they actually entail. When two accounts advance the same property, they may not actually be the *same* property. For example, Floridi and Sanders’ (2004) definition of Autonomy is satisfied if “the agent is able to change state without direct response to interaction” (p.26). For Hakli and Mäkelä (2019) however, this is only one part; an autonomous agent must *also* be able “to perceive the environment... [and] learn from experience” (p.264). Moreover, beyond ascriptions of AMS,

properties such as consciousness are “remarkably difficult to define and elucidate” (Gunkel, 2018a §3.3.2).

2.2 Thinking Otherwise

Avoiding properties, Relationism ascribes AMS “according to empirically observable, extrinsic relationships” (Gunkel, 2018a §6.1.2). Relationists reject the ontological pursuit of discerning morally relevant properties in advance of social interactions. Properties are, at most, mere *outcomes* of social interactions (Jecker, 2021) – “relations are ‘prior’ to the *relata*” (Coeckelbergh, 2012a p.45). In this sense, they are “*a posteriori* product of extrinsic social interactions” (Gunkel, 2018a §6.1.3). The intrinsic nature of an entity is therefore irrelevant for AMS – instead, we should focus on the *extrinsic* nature of our social relationships with a being (Coeckelbergh, 2010b). AMS is therefore grounded in how the *other* comes to supervene before us and how, “in the face of the *other*” (Gunkel, 2018a §6.1.2), we react.

In practice, the Relational approach reframes the debate of whether a robot *has* rights/AMS. Instead of asking the ontological question, Relationists ask an ethical one: whether a robot *ought* to have rights/AMS. This does not involve ascertaining what certain things *are* but discovering, “how we relate and respond to them in actual social situations and circumstances” (Gunkel, 2018a §6.2). How we should treat a being is therefore determined by our relationship with it (Nyholm, 2020). For example, our relationship with an animal fundamentally changes once we name, home, and pet it. We may come to consider it a “member of the family”, subsequently re-evaluating the degree of moral status we originally attributed it (Ibid p. 195). Gunkel (2018a §6.2.1) gives the example of Jibo, the world’s first family robot. It was advertised as not a thing, not a family member, but something between

the two. It is in this way, for the Relationist, that AMS is dependent on the role artificial beings play in our lives and the subsequent relationship we develop with them (Nyholm, 2020).

3. Relationist Anthropocentrism

3.1 Against Anthropocentrism

Prior to advancing Relationism, Gunkel outlines how past ethical theory has been unapologetically anthropocentric. Concerning personhood, he remarks that humans, who are already considered persons, “not only get to formulate the membership criteria of personhood but also nominate themselves as the deciding factor” (Gunkel, 2012 p.66). Anthropocentric definitions derive their criteria from human capacities whilst, simultaneously, nominating humans “both judge and jury” on who qualifies (Ibid p.67). Thus, they give unjust preferential treatment to humans over the interest of other beings (Haywood, 1997).

Property-based definitions of AMS often encapsulate top-down approaches. They identify a being *already* understood to have moral status and evaluate which of their properties should be necessary for *other* beings to also attain moral status. Significantly, the *original* being is, invariably, human. Sentience, Consciousness, and Autonomy are, for example, typically understood to be human traits. Consequently, these approaches utilise properties as proxies for human exceptionalism (Gunkel, 2012 p.67). The standard for AMS on such views is to *have* properties only humans are currently understood to have the capacity for. Such instances are clear examples of human chauvinism whereby they “specify relevant differences in ways that invariably favour human beings” (Haywood, 1997 p. 53). Moreover, it is then humans who are

to judge if these other beings meet such a standard. Conversely, for instance, humans would not accept a criterion that unjustly favoured artificial beings.

The Relationist, however, rejects any form of “moral centrism” (Mosakas, 2020 p.435). There are no pre-determined set criteria required for an entity to have AMS. Dependent on their individual social relations, Relationists decide on a case-by-case basis which other beings join their moral circle. Moreover, social relations are not exclusively attainable by humans – as Gunkel (2012) explains, “anything might take on a face” (p.179) and thus, we can form a *relationship* with anything. Consequently, the Relationist does not *directly* “decide if other beings are members of our club” (Coeckelbergh, 2012b p.236), they must adhere to the independent, extrinsic social context.

3.2 Accepting Anthropocentrism

Anthropocentrism is therefore typically understood as a criticism. For example, it is closely linked with the concept of Speciesism, whereby, for morally arbitrary reasons, a theorist gives preference to their own species at “the detriment of members of other species” (Haywood, 1997 p.52). Moreover, anthropocentric theories often give “purely instrumental consideration of non-humans” and, consequently, do not consider them “for their own sake” (ibid p.53). However, for those theorists that deny AMS ascription altogether, “this is exactly the point: robots are not humans” (Birhane and Dijk, 2020 p.209).

Joanna Bryson, for example, provides a purely instrumentalist account of robots (2010). Her thesis is that, as long as we avoid ascribing any morally ambiguous properties to robots, there is no harm in treating them purely as a means to an end. As Gunkel (2018a) elucidates: “No matter how capable they are, appear to be, or may become; we are obligated not to be obligated by them” (§4.0). A specific entity, *x*, cannot be dehumanised if that *x* is not

considered human (Bryson, 2010). Thus, Bryson embraces anthropocentrism by acknowledging that robots will be owned by humans, “created to be tools in our service” (Nyholm, 2020 p.190). By withholding moral consideration, when we inevitably instrumentalise robots, we cannot demean them. Consequently, Bryson (2010) states that “Robots should be slaves...*servants* you own” (p.110).

Gunkel, consistent with his anti-anthropocentric stance, raises multiple objections to Bryson’s instrumentalist position. Here, I present the strongest: that Bryson’s normative prescriptions “impose unrealistic...impractical restraints on behaviour” and “strain against lived experience...social norms and conventions” (Gunkel, 2018a §4.2.1). Firstly, it is unrealistic to impose practical restrictions on an AI having capacities for moral agency – eventually, on purpose or not, a system will breach convention (Ibid). Moreover, it contradicts real-world human-robot interaction. Not only do humans prefer working with autonomous robots over non-autonomous ones, they “empathise with what appeared to be robot suffering” (Ibid). A group of soldiers became so attached to their military robot “Boomer”, for example, they bestowed him two medals of honour and held a funeral for him (Garber, 2013). Humans cannot be expected to maintain a purely instrumental understanding of robots; through interaction, they form relationships, and, due to these relationships, “they will want to show moral consideration” (Nyholm, 2020 p.193).

However, do humans anthropomorphising robots imply that we should be open to ascribing them AMS? Birhane and Dijk (2020) “see no reason at all” (p.110). Social robots, on their view, are no different to any machine. Falling in love with a car, or “an espresso machine” (Ibid) they argue, does not imply we should grant them rights. We may treat these specific objects *better* than we otherwise would have but, fundamentally, we treat them well “as they were

the product of hard labour, expressions of human creativity...etc” (Ibid). By treating artefacts well, we are simply acknowledging their human makers.

In response, the Relationist would simply disagree we only treat artefacts well because a human had some part in their creation. The fundamental tenant of Relationism is that the *other's* intrinsic essence is irrelevant to the moral consideration it deserves. The Relationist would evaluate their social relationship between themselves and the artefact only after they have initially reacted to it. How it was made, what it was made of etc., is irrelevant to how the Relationist should treat it. Moreover, it is intuitive to our ordinary understanding of treatment toward artefacts. The soldiers, for example, did not hold a funeral and award two medals to Boomer because he was made/created by a human. They did it in virtue of their relationship with him: there are many robots of the same model, for instance, that would not receive the same treatment. *Who* created it, and *how* it was created, was irrelevant to how they treated it. Thus, the Relationist can remain committed to their anti-anthropocentric stance.

3.3 We the Relationist and the *Other*, the Robot

Despite Gunkel's anti-anthropocentric stance, Relationism is, itself, anthropocentric. Gunkel attempts to anticipate this by addressing the “unapologetically human” understanding of the *other* in the Levinasian philosophy Relationism follows (Gunkel, 2018a §6.3.1). Within Levinasian scholarship, he admits, the *other* exclusively refers to another human. Not only does ethics precede ontology for Levinas, but ethics itself is preceded by “a certain brand of humanism” (Ibid). Consequently, there can be no animal, or robot, *other*. Gunkel, therefore, adapts Levinas' philosophy “in excess of Levinas” (Ibid). However, he acknowledges that any re-articulation of Levinasian philosophy will still marginalise “any kind of technological *other*”

(Ibid). Accordingly, he distinguishes his Relationism from it. As aforementioned, for example, Gunkel (2012) himself maintains that “anything might take on a face” (p.179).

However, despite exceeding Levinasian human exceptionalism, Gunkel’s own Relationism remains anthropocentric. The Relationist’s issue is not who the *other can be*, it is who *they* are: necessarily human. Relationists ascribe AMS contingent on *their* independent, social-relational, phenomenological experience. That is, what is important when ascribing AMS is how *we* respond in the face of the *other*. Succinctly, following their denial of moral centrism, the Relationist assesses a being’s moral standing on a case-by-case basis. The upshot of this, however, is that it necessitates the significance of only one perspective: the Relationist’s. Relationist ascriptions of AMS require that *we*, the human, evaluate *our* extrinsic relationships with the *other*, the animal, robot, etc. Recall that Gunkel’s main challenge against anthropocentric theories is that, not only do humans set the qualifying criteria but that they are both “judge and jury”. Although the Relationist does not set a pre-determined qualifying criterion they, on a case-by-case basis, play both “judge and jury” on all claims of AMS. A robot’s morality is determined by its relationship with a human. It is then the human who reflects on whether this relationship is satisfactory. Although the Relationist does not directly decide what joins their moral circle - as they must adhere to extrinsic, independent social relations – it remains the Relationist who judges whether this relationship is viable. Moreover, it is the human who can invite the *other* into their home (Nyholm, 2020 p.195) and, as such, has the power to dictate their extrinsic relationship.

As Coeckelbergh (2012b) notes, Gunkel wishes us to avoid “anthropocentric domination” in deciding who has moral standing yet, at the same time, “thinks that it is up to us to decide” (p.236). The suggestion that humans have the capacity “to decide who is in and who is it out”

(Ibid) regarding AMS ascription, is inconsistent with Gunkel's anti-anthropocentric stance. Consequently, Gunkel's Relationism gives preferential treatment to humans over other beings (Haywood, 1997). As discussed, this need not be detrimental to the Relationist position. They could concede that their method of ascribing AMS *just is* anthropocentric; to uphold human welfare (Birhane and Dijk, 2020), they must remain "judge and jury." However, even if the Relationist does concede that its method AMS ascription is anthropocentric, it remains unable to secure sufficient action guidance.

4. Relationism's (lack-of) Action Guidance

The following section will critically evaluate the Relationist's ascription of AMS as an action-guiding theory. I will illustrate that it fails in this endeavour two-fold. §4.1 will illustrate, firstly, how it does not give a true account of moral status and, therefore, cannot ascribe universal normative prescriptions. §4.2 will argue that, regardless, the Relationist's method of ascribing AMS collapses into extreme meta-ethical relativism: as a result, it cannot guide our actions.

4.1 Relationism and Moral Status

Property-based approaches define their criteria for AMS ascription ontologically. Consequently, they provide a pre-determined guide on what deserves equal moral treatment. In this sense, their account prescribes to *every* moral agent. In contrast, the Relationist does "not truly offer an account of moral status, but only of particular agents' reasons vis-à-vis the individual at issue" (Jaworska and Tannenbaum, 2018 §5.5). Property-based approaches universally prescribe that a certain being should be treated well, whereas, for Relationists, this is dictated by our individual relationships. Fundamentally, for the Relationist, a being's moral status is dependent on it having a relationship with a human. Moreover, this ascription

of morality is only extended by that *particular* human. Without this relationship, there would be no reason “not to kill the being” (Ibid). A true account of moral status, however, “should give every moral agent...reasons to protect that being” (Ibid). Relationism cannot secure this as AMS is ascribed on an individual, case-by-case basis. Consequently, they cannot give any universal normative prescriptions. On what grounds, for example, could a Relationist condemn the ill-treatment of their family robot, Jibo, by someone who has no relation to it or them?

Mosakas (2020) has suggested that a Relationist may contend this is all moral status is. That it is a “construct,” merely encapsulating “different agents with their own subjective reasons and dispositions that determine how they act toward others” (p.434). This suggestion, however, is inconsistent with the Relationist’s aim. Gunkel’s re-articulation of the AMS debate, for example, specifically distinguishes Relationism from this interpretation. As aforementioned, Gunkel (2018a) wishes to avoid asking whether robots *have* rights/AMS and instead asks if they *ought* to have rights/AMS (§6.2). In this sense, his rejection of the ontological and focus on the ethical, necessitates a case-by-case approach. However, this does not, therefore, mean that for the Relationist moral status is merely a construct: simply that their answer only pertains to specific beings. Regardless, Relationism remains too weak to bear universal normative implications. The Relationist’s case-by-case ascription of AMS, although remaining committed to *genuine* moral status, can only account for individual accounts, not full.

Following Tollon and Naidoo (2021), this issue pertains to the Relationist’s fundamental case-by-case method of ascription as, as individuals “we often misinterpret the moral nature of relations” (§6). Although their love for books compels *them* to treat books with respect, it

does not, therefore, follow that books *objectively* deserve to be treated with respect – let alone mean that “books do in fact have moral status” (Ibid). As Birhane and Dijk (2020) argued, treating something well does not mean that even we ourselves think something *deserves* to be treated well. Our individual relations are just that: individual. Moreover, they are subjective and independent. Thus, on Relational accounts, whereby whatever we deem our relation to be with an artefact determines their moral character (Tollon and Naidoo, 2021), this ascription only extends to *ourselves*. On a Relationist account, we are unable to secure objective morality.

Recall Boomer, the military robot. The soldier’s treatment of Boomer, according to Relationism, was in virtue of their relationship with him – no doubt influenced by the countless lives Boomer saved (Garber, 2013). Those who did not have this relationship with Boomer, however, nor knew anything about his heroics, would not share the same sentiment. “‘He’ is one of many just like ‘him’” they might jest. Significantly, the soldiers’ subjective, independent ascription of moral character to Boomer, is specific to them. Whatever moral character they’d ascribe Boomer, it would not exceed their specific social-relational context.

4.2 Relationist Relativism

Given that Relationism only accounts for individual ascriptions of AMS, it cannot secure objective moral standards. As a result, it collapses into extreme meta-ethical relativism. That is, “the view that no moral judgments or standards (about any moral questions) are objectively true (or, correct) or false (or, incorrect)” (Moser and Carson, 2001 p.3). For example, each Relationist will have their own subjectively independent perception of a specific entity. Consequently, they will form different relationships and, therefore, ascribe

differing prescriptions of treatment. Relationism cannot dictate which prescriptions are *correct* and should be followed objectively.

As a result, Relationism cannot guide our actions. Without “central moral properties or guiding principles”, we get “bogged down in a sea of relative judgements” (Mosakas, 2020 p.434). On an individual, case-by-case basis, it is our independent relationships that guide action for the Relationist. However, when two individuals’ relationships with the same entity conflict, Relationism is unable to guide which is “true (or, correct) or false (or, incorrect)” (Moser and Carson, 2001 p.3). Alternatively, for property-based approaches, their pre-determined criteria are also their objective truths and falsities. Simply, a being has AMS if it has the necessary property that a specific theory requires. Property-based approaches are therefore disjunctive insofar that they either ascribe a being AMS or they do not. This allows for intuitive action guidance; if a being has AMS then we should treat it as such and if it does not, then we should not. The Relationist, however, allows cases whereby on one distinct account a being *should* be ascribed AMS whereas simultaneously, on another, the being should *not* be ascribed AMS. For Coeckelbergh, as Tollon and Naidoo (2021) remark, it is simply a case of an entities’ appearance provoking “the correct emotional response” (§6). What happens when one Relationists’ emotional response is of the *correct* kind, whereas another’s is *wrong*? Both of these ascriptions would be *true* according to Relationism as, as Müller (2021) observes, “anything goes” (§Nasty Reconstruction (2)). Relationism does not tell us which ascription should be followed. Consequently, the Relationist’s method of AMS ascription cannot guide our moral decisions.

4.2.1 Gunkel's Ethical Pluralism

Gunkel (2018a) attempts to embrace relativism by adopting a kind of pluralism. Giving the example of the term in physics, he points out that being relative “need not be construed negatively” (Ibid, §6.3.2). We can remain critical of *moral* relativism, he argues, whilst acknowledging that moral standards “are socially constructed formations that are subject to and the subject of difference” (Ibid). Accordingly, he adopts ethical pluralism, explaining how it can be “a middle-ground between the extreme version of relativism and moral absolutism” (Mosakas, 2020 p.435). However, Gunkel neglects to apply it to Relationism. The most he does, as Mosakas points out, is give “a number of citations” (Ibid).

This is significant for two reasons. Firstly, ethical pluralism is typically understood to entail “a plurality of basic moral values” (Mosakas, 2020 p.435). Recall, however, that Gunkel is highly critical of any centrality of moral values/properties. It is a fundamental tenant of Relationism that there are no pre-determined criteria: prior to relations, there is no moral guidance. In contrast, the Pluralist accepts a whole host of irreducible criteria. Gunkel needs to explain how there can be a middle ground.

Secondly, Gunkel does not explain how pluralism allows Relationism to eschew the charge of relativism. By adopting pluralism, Gunkel is supposing that we are now able to make decisions about whether one Relationist's ascription is better than another's. Yet, not only does he fail to give an account of how Relationism is now able to do this, even if he did, we would still lack a guiding principle regarding which standards are “more authoritative than others” (Mosakas, 2020 p.435). Moreover, if Gunkel did provide a guiding criterion, he would commit to a form of moral centrism. Without one though, his view again collapses into moral relativism. As

Tollon and Naidoo (2021) comment, we can accept that moral standards change over time, but not when such change is dictated by “epistemically frail terms” (§6).

5. Conclusion

In sum, I began this paper by exploring the Relationist’s critical stance toward property-based approaches of ascribing AMS. Despite largely agreeing with the critiques raised, I identified two condemning criticisms of the Relationist’s own, social-relational case-by-case approach. Firstly, I illustrated how, despite their anti-anthropocentric stance, the Relationist’s method of individual case-by-case AMS ascription is itself anthropocentric. Secondly, my analysis of their case-by-case method demonstrated that the Relationist can neither secure objective morality nor provide sufficient action guidance. Despite considering Gunkel’s ethical pluralism (2018a) I, following Tollon and Naidoo (2021) and Mosakas (2020), concluded that this is not a satisfactory response. Thus, this essay has illustrated that not only is the Relationist’s method of ascribing AMS anthropocentric, it cannot secure sufficient action guidance.

Reference list:

- Andreotta, A. (2020). The hard problem of AI rights. *AI & Society*, 36 (1). [Online]. Available at: doi:10.1007/s00146-020-00997-x. [Accessed 16 December 2021].
- Birhane, A. and Dijk, J. van. (2020). Robot rights?: Let's talk about human welfare instead. In: *AIES '20*. 2020. Association for Computing Machinery (ACM). [Online]. Available at: doi:10.1145/3375627.3375855. [Accessed 16 December 2021].
- Bryson, J.J. (2010). Robots Should Be Slaves. In Wilks, Yorick. (Ed). *Close Engagements with Artificial Companions: Key social, psychological, ethical and design issues*. Philadelphia, PA: John Benjamins, pp. 63-74.
- Coeckelbergh, M. (2010a). Moral Appearances: Emotions, Robots, and Human Morality. *Ethics and information technology*, 12 (3). [Online]. Available at: doi:10.1007/s10676-010-9221-y. [Accessed 16 December 2021].
- Coeckelbergh, M. (2010b). Robot Rights? Towards a Social-Relational Justification of Moral Consideration. *Ethics and information technology*, 12 (3). [Online]. Available at: doi:10.1007/s10676-010-9235-5. [Accessed 16 December 2021].
- Coeckelbergh, M. (2012a). *Growing Moral Relations*. London: Palgrave Macmillan.
- Coeckelbergh, M. (2012b). David J. Gunkel: The machine question: critical perspectives on AI, robots, and ethics. *Ethics and Information Technology*, 15. [Online]. Available at: doi.org/10.1007/s10676-012-9305-y. [Accessed 16 December 2021].
- Danaher, J. (2020). Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism. *Science and engineering ethics*, 26 (4). [Online]. Available at: doi:10.1007/s11948-019-00119-x. [Accessed 16 December 2021].
- Danaher, J. (2021). What Matters for Moral Status: Behavioral or Cognitive Equivalence? *Cambridge quarterly of healthcare ethics*, 30 (3). [Online]. Available at: doi:10.1017/S0963180120001024. [Accessed 16 December 2021].
- Eskens, R. (2017). Is Sex With Robots Rape? *Journal of Practical Ethics*, 5 (2). [Online]. Available at: <https://www.ipe.ox.ac.uk/papers/is-sex-with-robots-rape/> [Accessed 16 December 2021].
- Floridi, L. and Sanders, J. (2004). On the Morality of Artificial Agents. *Minds and machines*, 14 (3). [Online]. Available at: doi:10.1023/B:MIND.0000035461.63578.9d. [Accessed 16 December 2021].
- Garber, M. (2013). *Funerals for Fallen Robots*. [Online]. The Atlantic. Last updated: 20 September 2013. Available at: <https://www.theatlantic.com/technology/archive/2013/09/funerals-for-fallen-robots/279861/> [Accessed 16 December 2021].

Gunkel, D. J. (2012). *The machine question: critical perspectives on AI, robots, and ethics*. Cambridge, Mass: MIT Press.

Gunkel, D. J. (2018a). *Robot rights*. Cambridge, Mass: MIT Press.

Gunkel, D. J. (2018b). The other question: can and should robots have rights? *Ethics and information technology*, 20 (2). [Online]. Available at: doi:10.1007/s10676-017-9442-4. [Accessed 16 December 2021].

Hakli, R. and Mäkelä, P. (2019). Moral Responsibility of Robots and Hybrid Agents. *The Monist*, 102 (2). [Online]. Available at: doi:10.1093/monist/onz009. [Accessed 16 December 2021].

Hayward, T. (1997). Anthropocentrism: A Misunderstood Problem. *Environmental Values*, 6 (1). [Online]. Available at: doi.org/10.3197/096327197776679185. [Accessed 16 December 2021].

Jaworska A, Tannenbaum J. (2018). *The grounds of moral status*. [Online]. Stanford Encyclopedia of Philosophy. Last updated: 10 January 2018. Available at: <https://plato.stanford.edu/archives/spr2018/entries/grounds-moral-status/> [Accessed 16 December 2021].

Jecker, N. S. (2021). Can we wrong a robot? *AI & Society*. [Online]. Available at: doi:10.1007/s00146-021-01278-x. [Accessed 16 December 2021].

Mosakas, K. (2020). On the moral status of social robots: considering the consciousness criterion. *AI & Society*, 36 (2). [Online]. Available at: doi:10.1007/s00146-020-01002-1. [Accessed 16 December 2021].

Moser, P and Carson, T. (2001). *Moral relativism: a reader*. Oxford: Oxford University Press.

Müller, V. (2021). Is it time for robot rights? Moral status in artificial entities. *Ethics and Information Technology*, 23. [Online]. Available at: doi.org/10.1007/s10676-021-09596-w. [Accessed 16 December 2021].

Nyholm, S. (2020). *Humans and Robots: Ethics, Agency, and Anthropomorphism*. Washington, DC: Rowman & Littlefield International.

Singer, P. (1981). *The expanding circle : ethics and sociobiology*. Oxford: Clarendon Press.

Sullins, John P. (2006). When is a robot a moral agent. *International Review of Information Ethics*, 6 (12). [Online]. Available at: <https://scholarworks.calstate.edu/downloads/fj2362670?locale=en> [Accessed 16 December 2021].

Tollon, F and Naidoo, K. (2021). On and beyond artifacts in moral relations: accounting for power and violence in Coeckelbergh's social relationism. *AI & Society*. [Online]. Available at: doi:10.1007/s00146-021-01303-z. [Accessed 16 December 2021].

Véliz, C. (2021). Moral zombies: why algorithms are not moral agents. *AI & Society*, 36 (2). [Online]. Available at: [doi:10.1007/s00146-021-01189-x](https://doi.org/10.1007/s00146-021-01189-x). [Accessed 16 December 2021].