



On and beyond artifacts in moral relations: accounting for power and violence in Coeckelbergh’s social relationism

Fabio Tollon¹ · Kiasha Naidoo²

Received: 12 July 2021 / Accepted: 7 October 2021 / Published online: 17 October 2021
© The Author(s) 2021

Abstract

The ubiquity of technology in our lives and its culmination in artificial intelligence raises questions about its role in our moral considerations. In this paper, we address a moral concern in relation to technological systems given their deep integration in our lives. Coeckelbergh develops a social-relational account, suggesting that it can point us toward a dynamic, historicised evaluation of moral concern. While agreeing with Coeckelbergh’s move away from grounding moral concern in the ontological properties of entities, we suggest that it problematically upholds moral relativism. We suggest that the role of power, as described by Arendt and Foucault, is significant in social relations and as curating moral possibilities. This produces a clearer picture of the relations at hand and opens up the possibility that relations may be deemed violent. Violence as such gives us some way of evaluating the morality of a social relation, moving away from Coeckelbergh’s seeming relativism while retaining his emphasis on social–historical moral precedent.

Keywords Moral status · Relational turn · Foucault · Power · Moral consideration

1 Grounding moral status

Recent research in the ethics of AI (Artificial Intelligence) has come to focus on the grounds for moral status (see Coeckelbergh 2010, 2012; Gunkel 2012; Torrance 2014; Danaher 2017a, b; Tollon 2020). What is it about entities that makes them worthy of moral concern? Who (or what) should be accorded rights? These are some of the key questions in moral philosophy. Under this set of questions are concerns regarding the rights of various entities. We talk about human rights, women’s rights, animal rights, and now even “rights of nature”, that is, the rights we might accord to non-sentient entities (such as Whanganui river in New Zealand) (Gellers 2021: 2). It is still an open question whether it is “time” for robot rights, or whether talk of robot rights

is at all coherent, and so we bracket this question in our discussion (Müller 2021). In this paper, we are interested in moral concern more generally, and although this account surely has significance to scholars who are working in the field of “robot rights”, we do not directly engage with the arguments in that field (Gunkel 2014, 2018; Gellers 2021). More specifically, our attention is directed towards the ways in which moral status might be grounded.

It was first thought, in the Western philosophical canon, that only persons of European descent were worthy of moral concern. This was then expanded (not without much resistance) to include all members of *Homo sapiens*. There is now a general consensus that non-human entities may be subjects of moral concern, and the animal rights movement (and its success) is a testament to this. Moral concern here generally means that we should respect these entities, that harming them would be wrong. What then about our artificial companions? Does the rise of sophisticated and interactive AI mean that we need to reconsider our criteria for moral status, extending it, for the first time, beyond the consideration of biological entities? In the past, it has often been the case that moral concern was unjustifiably denied, and some of the worst acts in human history were based on false claims about who or what is worthy of moral concern. It matters that we get this right.

✉ Fabio Tollon
fabiotollon@gmail.com

Kiasha Naidoo
kiasha.naidoo@gmail.com

¹ Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany

² Centre for Humanities Research and Department of Philosophy, University of the Western Cape, Cape Town, South Africa

One of the key debates in the grounding of moral concern, as this relates to AI is between “properties” approaches and “social relational” approaches. Instead, we choose to call the properties approach a “realist” approach to moral status. This is not to say that the social relational account is any less “real” (in the sense of actually existing in the world), but rather to emphasize the role that human perception plays in each account. For a realist, the moral status of an entity is formulated as independent, or prior to the perception of the relevant entities. For the social relationist, however, perception matters a great deal.

In this paper, we will first outline and critique the realist account of moral status. Following this we provide an exposition of social-relational accounts while also showing that realist accounts face a variety of challenges in terms of their ability to account for the ways social and political life come to inform who or what is deserving of moral concern. Subsequently, we outline a social-relational account of moral concern, which does not fall victim to the criticisms levelled against realist accounts. However, this is not to say that such positions are not without shortcomings. We will specifically discuss the relativistic implications of Coeckelbergh’s (2010) social relational account and seek to remedy this. We do this by incorporating a Foucauldian–Arendtian notion of power and violence into the conception of social-relationism, which both contributes to its explanatory potential and helps such accounts avoid the charge of relativism.

2 The real and the relational

For the realist, there are experience independent answers to questions regarding moral status. For example, if we were to ask whether *X* is conscious, suffering, worthy of moral concern, etc., the realist would respond that there are objectively correct answers to such questions, in the sense that such answers would tell us what is really going on, and are perception independent (Torrance 2014, p. 11). For the realist, there is a “fact of the matter”, in descriptive terms, for our grounding of moral status. There are certain measurable properties that determine whether an agent is worthy of moral concern or not. For a social-relationist, however, there are no “intrinsic” properties of agents that make them worthy of moral concern. Rather, such concern is grounded in how these entities appear to us (Coeckelbergh 2010, p. 214). This concern is not “out there”, but rather is to be found in our relation to entities, and the social norms that shape (and co-determine) such relations. Another way to understand our argument is that merely descriptive accounts of entities are not enough to make a proper determination of their moral status. That is not to say that such descriptive criteria are useless, but rather that they should not exhaust our analysis.

For this reason, we delve into more normative metaethical terrain to support a social-relational view of moral status.

3 Realism

The realist claims that there are objective answers to our grounding of moral status. On this view, questions regarding our treatment of potential holders of moral consideration are determined by external factors. For example, those related to how we go about ascertaining whether an entity is phenomenally conscious or not (Torrance 2014, p. 13). These factors are external in the sense that their verification lies outside the domain of moral philosophy as such, and so appeals to philosophy of mind, for example, are commonplace. The link between being conscious in this sense and being a carrier of moral value, according to Torrance, is that consciousness comes with an associated experiential valence (2014, p. 14). As humans, our experiences are often coloured in different ways. From the irritating noise of the air conditioning system, to the pleasant white-noise of background conversation in a coffee shop, each is valenced according to their experiential content, and can therefore contribute to our ill- or well-being (Torrance 2014, p. 14). The varying degrees of arousal, whether negative or positive, determine the level of suffering or satisfaction for a given conscious creature. In this sense, there are strong biological features which give us the capacity to experience things as favourable or not, and these capacities are broadly independent of any social criteria. Of course, this is not to say that social criteria play no role at all, but rather that non-social criteria are the bedrock upon which moral concern rests. While societal values may change when and how we come to embrace new others as moral patients, this does not change the fact of the matter that all such entities must be conscious (Torrance 2008, 2014). The realist is committed to grounding moral status in some objective property of the agent in question, where this property is independent of our experience of it. We have focussed on phenomenal consciousness, but in principle any approach that seeks to ground moral status in some specific ontological property of the entity would be considered “realist” on our account. This has important implications, as we will make clear in what follows.

4 Problems with realism: the slippery slope(s)

The first issue, which we term the problem of indirect access, concerns the relationship between our intuitions and our best science. The realist claims that sentience, or some kind of consciousness, is essential for entering the moral circle. However, we do not have direct access to such

states (except for perhaps in our own case). We do, however, have indirect access. Such indirect access may permit us, based on various criteria, to reliably infer whether an entity is sentient or not. Take the example of fish cognition. It is commonly supposed that fish are not sentient in the way that various mammals are. However, fish display an array of cognitive skills that, should a mammal display them, we would happily report as evidence of their sentience (Brown 2015). The evidence is mounting that our neglect of fish well-being is of serious moral concern (Brown 2015; Tollon 2020). Their behaviours, however, do not trigger in us the same moral intuitions that are prompted when we view other mammals. The issue with the realist position, therefore, is that it assumes that we can, in fact, have direct access to the seemingly “objective” criteria underpinning moral status. However, it does not seem that we have access to these criteria, and, moreover, when we investigate these criteria, we seem to bring our own biases to the table.

The second issue that realism faces is the moral progress problem. It is broadly metaphysical and relates to the first issue above: how can such an account accommodate the fact that who or what comes to be worthy of moral concern has, in fact, changed over time. If we have objective answers to questions regarding a moral concern, then it seems we would have to admit that slaves in the sixteenth century were moral subjects. Objectively speaking, enslaved people possessed all the criteria required for the realist to grant them moral status. They were persons sharing all the necessary criteria for them to be considered moral subjects. In practice, however, this is not how things played out (Césaire and Kelley 2000). The practice of slavery was politically encoded, and thus the criteria of moral status were legitimated through political and social practices. No matter that the conditions were met for slaves to be moral subjects on the realist account, they were not treated as such. The realist position does not allow for us to fully acknowledge the way in which those in positions of influence may abuse the claim to objective moral grounding as a justification for what seem like morally questionable actions. This is particularly so if the grounding is based on the capacity for individual consciousness—the identification of which is often subject to our morally fallible intuitions, as noted above.

Our third criticism of the realist approach, what we will call the entity grounding problem, stems from an argument made by Coeckelbergh which focuses on ontological properties (2010). The realist approach (which seems *prima facie* to be broadly utilitarian with its focus on affective capacities), seeks to justify a basis for moral concern grounded in the ontological properties of individual entities. Problems with these views include seeming exceptions to these criteria, the difficulty with marginal cases and practical difficulties as discussed above. A response to these concerns is found in the move toward a social-relational approach (the

“relational turn”) which seeks to ground moral concern in a way which may better account for the morally relevant relations between entities and the contexts in which they exist. Our previous suggestion that the realist approach may be used to justify (and codify) practices which we now deem immoral suggests that entities are inextricably intertwined with their social and relational contexts. This involvement has an unavoidable bearing on moral prescriptions, norms, and practices. As we will seek to emphasise, not only are moral considerations partly a function of social and relational factors, the composition of the environment is also shaped by moral and political influences and in turn shapes the set of moral possibilities. In what follows we will outline Coeckelbergh’s account and show how it plugs the gaps exposed by the realist account presented above.

5 Social-relational moral concern

The most developed social-relational account of moral concern is due to Coeckelbergh (2010), where he argues against both direct (utilitarian, deontological) and indirect (virtue ethical) criteria for moral standing. He argues that the aforementioned all posit some essential ontological property, and these properties (such as sentience), face impossible issues of justification. In place of these, Coeckelbergh suggests that we look toward the concrete way in which entities are experienced by us, in other words, how they appear to us (2010, p. 214). Their appearance, in conjunction with how we decide to respond, is the grounding principle of relational accounts. This approach shifts the burden of proof away from the need to investigate which ontological property is the “correct” one, and instead focuses on the social network in which the entity is embedded. Social-relationism is therefore both context- and subject-dependant (Coeckelbergh 2010, p. 214). In what follows we will show the social-relational account can address the three issues with realism outlined in preceding sections.

Due to the fact that the social-relational account allows for expansion of our moral circle over time, it avoids the problems of indirect access and of moral progress, and directly responds to the entity grounding problem. The changing criteria of moral consideration in different historical contexts can easily be accommodated, as the relational account makes no *a priori* claims regarding just what these criteria ought to be. It can therefore accommodate changes in our scientific understanding of what the “correct” criteria may or may not be for our ascriptions of moral patiency.

Instead of insisting upon (at best) difficult-to-attain epistemic standards regarding the moral natures inhering in objects and subjects, it shifts focus to the relationship between subject and object. In this way, the theory makes room for our grounds of moral concern to change over time.

As noted earlier, the expansion of our moral circle did not occur instantaneously: it took hundreds of years and countless struggle movements for us to get to the emancipation of slaves, women's suffrage, and the liberation of animals (Singer 1975). These changes were not brought about because we suddenly realised that all persons are in fact sentient. If any biological entity is now deemed to be sentient, it seems obvious that they were also sentient before this designation. Therefore, there is a strong sense in which we make these judgments not on the basis of essential ontological features, but rather on the basis of how the entities appear to us, in a given social and cultural context. The emancipation of slaves, the enfranchisement of woman, and the liberation of animals are thus all different ways in which attitudes in society have co-shaped who or what comes to be worthy of moral concern. None of this necessarily turns on whether any of the aforementioned are sentient or not (in an ontologically demanding sense). In this way, the social-relational approach can meet the entity grounding critique of the realist position above. This argument, therefore, invites us to consider that the emergence of artificially intelligent systems, whether sentient or not, will have some kind of relation to us, and will invite certain kinds of behaviours and attitudes. We will treat them in certain ways, and we will not be the same individuals we were before their introduction to society (Coeckelbergh 2010, p. 215).

While the social-relational approach above may solve issues faced by the realist the charge of moral relativism is one that must be met. In our view, Coeckelbergh's position, as it stands, does not fully escape this charge. This is because it does not do enough to address how power relations are concurrent in social relations in any given society. In what follows we will show how the social-relational account, as it stands, cannot successfully meet the charge of relativism, and we will subsequently extend Coeckelbergh's account and show how it can accommodate such power relations, giving it greater explanatory power and philosophical depth.

6 Social-relationalism and relativism

The charge of relativism rears its head due to the fact the social-relational approach grounds moral status in the way entities appear to us. This epistemic point implies that should such an "appearance" provoke the correct emotional response, then the entity in question would be a moral patient. This is troubling for two reasons. First, consider how we often misinterpret the moral nature of relations (due to our unfounded attribution of psychological characteristics to clearly inanimate objects). For example, the authors of this paper both have a shared love for physical books. We, therefore, expect ourselves and others not to cause undue damage to books. People ought to respect books, and therefore

have a certain responsibility towards them. However, this does not mean that books do in fact have moral status. Just because we care about books and consider mistreating them problematic, does not follow that books have moral status. On the relational account, our perception (and subsequent response) to artifacts is what determines the moral character of that artifact. Thus it seems to be a case of "anything goes", where whatever we care about (in some trivial sense) becomes a moral subject (Müller 2021: 5).

The second point is that it seems inappropriate to ground moral status in appearances alone: this could result in moral descriptions subject to the limits of the subjective experience and thus is blind to aspects of moral relations which are not immediately available or harder to detect. Although we might be inclined to focus on clear moral relations, for example as manifesting in the sorts of actions which seem to straightforwardly arise out of a set of moral relations, such as in cases of murder and other clear violations of individual autonomy. However, that these are clear instances of moral relationality does not mean that all instances of moral relationality will be as clear. Some might be much less readily visible.

While it seems realist approaches fail because of their stringent search for ontological properties, social-relational accounts fail in the opposite direction in their focus on only appearances. There are those, however, who seem to embrace this kind of relativism. David Gunkel, for example, argues that,

'one can remain critical of "moral relativism," in the usual dogmatic sense of the phrase, while being open and receptive to the fact that moral standards—like many social conventions and legal statutes—are socially constructed formations that are subject to and the subject of difference' (2018).

We agree with Gunkel that our moral standards and practices do in fact change over time, and that this fact should significantly inform our grounding of moral status. Gunkel proposes pluralism without relativism, which entails an acknowledgement that moral truths are not 'absolute', but nor does 'anything go' (Gunkel 2018, p. 181). These claims are broadly consistent with the arguments we have advanced up until now. However, Gunkel does not successfully provide an account of how exactly relational accounts are meant to escape relativism. Even by acknowledging a kind of pluralism, we would still need a way to navigate our moral landscape. Genuine relativism does not allow for one option to be better than any other, and the introduction of pluralism is meant to solve this problem. However, on Gunkel's account, much is still missing, as he, such as Coeckelbergh, does not address the epistemic issues raised by the usage of appearance to ground moral concern. Basically, to remain critical of moral relativism, we must avoid

it. Accepting that moral standards can and do change over time is not necessarily relativistic, but when such change is accounted for in epistemically frail terms, the relativistic critique can pack a punch. Specifically, the issue remains that if we are to ground moral relations in how entities appear to us (and our responses to this appearance), we risk making the same mistakes that Gunkel seeks to avoid. Namely, by not investigating what lies behind these appearances we will not be able to get a handle on what exactly pluralism without relativism is.

7 Toward a view of power in social-relationality

As discussed, Coeckelbergh's (2010) account of a relational moral theory leads to a kind of relativism and insufficiently avoids a re-inscription of prevailing moral norms in its privileging of 'apparent' moral relations as a determinant for moral standards. This view of moral concern as based on social-relational dynamics, nevertheless, holds weight and can be developed to avoid relativism and cater to Gunkel's (2018) desire for a framework which comprehends the dynamic nature of moral standards over time and across (cultural, religious, ethnographic, etc.) spaces. The dynamism of moral values, understood as relational norms, urge us to consider the fundamental normative structures which contribute to the curation of moral relations and possibilities. That is, if there is a set of moral possibilities, there is also its production. The Foucauldian notion of power/knowledge allows us to consider social relations as defined by the power which comes to inform what counts as valid knowledge under a given discursive structure, including moral 'knowledge,' in a society (Foucault 1978). As indicated by the term power/knowledge, a Foucauldian perspective opens up the opportunity to consider whether the pervasive role of power can explain moral pictures with a greater degree of nuance than realism and Coeckelbergh's social relationism. By conceptualising power at the hand of both Michel Foucault and Hannah Arendt, we draw attention to the profound role of power in curating social relations, allowing a move away from relativism while maintaining a social-relational conception of moral possibilities.

The extent to which power is involved in colouring social relations is one which Coeckelbergh fails to substantively acknowledge. His preliminary remarks regarding the relation between trust, power and information technology (Coeckelbergh 2011) include a consideration of power as adding a vertical dimension to normative frameworks, conceiving of the state as an entity with the capacity to impose influence/power in a downward fashion. This understanding of power relations as orthogonal to moral relations arises from

a view of power which is strictly top-down, the problems with which will be expounded upon below.

In the article we have focused on, Coeckelbergh discusses social-relationality in its potential application to political affairs, which could purportedly cut through the obfuscation of bureaucratic legal-governmental structures and illuminate the true nature of relational dynamics (Coeckelbergh 2010). That is, a consideration of social-relationality can access information about relations which, due to powerful institutions, would otherwise be imperceptible. In only seeing how social-relationality may be applied to situations of power, it is not recognised as inherent to these social relations and thus not seen as directly curating moral concern. Coeckelbergh's (2015) argument, however, regarding the master–slave relation as descriptive of human–technology and human–human relations seems to imply the inherency of power in moral relations rather than as a feature which can be considered on top of/in addition to them. The exact nature of this inherency nevertheless requires explanation and the problem of moral relativism in social-relationism remains unsolved.

If power is conceived as relational and intersubjective, as Coeckelbergh (2010) understands the moral concern, we come to see that power is not merely orthogonal to moral relations nor are they parallel—they are intertwined, always going together.

8 A conception of power

The conception of power we will use to understand moral relationism is distinct from the "realist"¹ (hereafter, the "conventional") understanding used frequently in political science due to its understanding of power as top-down, organised power which is seen as coercive and identified with the means of coercion. That is, to have power is to be in possession of the means to exercise coercive tactics. This conventional understanding posits violence and domination as inherent to political engagements (Breen 2007, p. 352), defining the political in terms of power. This understanding is limited in its explanatory use, particularly in regard to relations outside of the political realm. We are particularly interested in the possibility that power is not possessed by an individual or entity but is relational, as described in Foucault's writings on power (Foucault 2001). A conception of power, formulated by Jacob Maze (2018) synthesizes Hannah Arendt and Michel Foucault's philosophical work on power, offering a picture of power which comes to explain relations more generally. Arendt and Foucault's theories,

¹ Power as understood in the realist paradigm bears no resemblance to moral realism.

although substantively opposing on multiple counts,² converge (Allen 2002; Maze 2018) on a number of specific points regarding an understanding of power and its relation to the idea of violence. This convergence allows us to describe the pervasive nature of power while not rendering the analysis of power meaningless. That is, we can describe power as widespread while the idea of violence affords a moral dimension. This helps to avoid the argumentative slippage which might lead to the conclusion that since every relation is one of power and relations are unavoidable, it is impossible to attribute a moral character to these relations. This all rests on a careful conceptualisation of power.

Power is at base understood as strategic action ‘employed in an attempt to control a situation’ (Maze 2018). In a similar vein, Arendt (1970, p. 36) draws on Voltaire’s theorisation of power, defined as the exercised control of the actions of others in accordance with one’s own will. This can be expressed simply as coercion. Arendt and Foucault converge on the point that power, although understood as coercive, is only present if there is the potential for those being coerced to resist a given attempt to control. The absence of the potential to resist renders the relation devoid of power, instead indicating that the situation bears a greater resemblance to violence rather than an exercise of power.

The basic definition of power as coercive leaves room for the possibility that it is not a term reserved for describing a top-down means of coercion. Rather, it allows for the Foucauldian view which holds that power is exercised not only in conventional political relations such as that between governments and their citizens but also in a much broader set of relations including between citizens, firms and citizens and, as we are exploring, between individuals and artifacts. Arendt and Foucault concur on the supposition that power is not a property or quality possessed by entities such as organisations or the state (Breen 2007). It cannot be understood as that which inheres in an entity itself since power is something which depends upon there being a subject of any intended coercion. Arendt (1998, p. 200) states, ‘power springs up between men when they act together and vanishes the moment they disperse,’ implying its intersubjectivity and thus its relationality. The move to construct a view of power as such parallels Coeckelbergh’s move away from grounding moral concern in the properties of entities. That is, power and moral concern are similar in kind and neither are had by one entity over another as a result of some set of properties but rather, it describes what is going on between these entities.

² For example on the distinction between the private and public. While Arendt sought to maintain the distinction, Foucault questioned this presupposition. While we do not address this distinction here, we proceed a Foucauldian frame, questioning the possibility of a “coercion-free” relations (Villa, 1992:712).

9 Power and social relationism

If power is understood as an attempt to control, then we can in the first instance assess instances of social-relationality to the extent that they may constitute such an attempt. Social relations are coloured by the degree to which power is present and this presence can be considered in specific relational instances as well as in preceding instances and possibly succeeding where relevant. This means simply that exercises of power—attempts to control—including past and potential attempts, shape prevailing social relations and the future of these relations. The importance of understanding social–historical dynamics for informing current moral concern is noted by Coeckelbergh (2010, p. 214). As discussed, however, Coeckelbergh fails to theorise an understanding of moral relations which necessarily involves the role of power. The relation between husbands and wives, for example, in conventional nuclear family setups is interesting precisely because of the possibility that it may be characterised by an attempt to control. The issues of domestic violence and the sharing of duties in the household are coloured by dynamics of patriarchal power structures, dependency, coercion and perhaps even a blindness to roles of power, which is one way in which the possibility to resist may be unavailable. Such relations are curated by historical relations (often) of power. It should become clear that many instances of moral relation between entities, are distinctly characterised by attempts to control. If we are considering rightness or wrongness, we are already also considering the extent to which coercion may be occurring and whether this relation is violent, both in the everyday use of the term (as the perpetration of physical abuse) and understood as a severe restriction of/abolition of the capacity to resist coercion. In other words, the relations in which power functions (or has shaped) also tend to be those which we want to consider from a moral point of view.

Social-relationality advocates for a picture of moral concern which, as discussed in our critique of Coeckelbergh’s theorisation, arises from how relations appear rather than how they actually are. This may lead to a myopic view of power because strategic action to control is not always apparent, and if it is, this is seldom to its fullest extent. The frequent inability to clearly identify power motivates a plethora of research, for example, in social and political philosophy, political science, and advocacy work, with the explicit aim of uncovering the nature and degree to which power shapes relations. This concern extends to the moral consideration of technological application(s). Examples from this realm of considerations can help to understand the role of power and potential for violence in moral considerations. Such research touches on, for example, the ways in which algorithmic systems are developed to control behaviour in some or other way. This could include, for example, YouTube’s

recommender algorithm which has been found to promote conspiratorial content (Burr et al. 2018; Alfano et al. 2020; Tollon 2021). It is this attempt to control behaviour—an exercise of power—which animates a discussion of the (im)permissibility of such applications.

It is in this sense that technology appears deeply involved with moral agents and in situations where moral (or immoral) actions may be performed. This serves to broaden the scope of the set of actions we deem worthy of moral evaluation, and the kinds of things that are relevant in moral situations (Peterson and Spahn 2011). The potential for coercive relations in technology is what makes such relations subject to consideration in the realm of moral philosophy. This goes along with our view that technology is not a neutral tool in our environments which merely slots into the moral nature of a decision, but also a producer of the moral possibilities which inform these decisions (Johnson 2006; van de Poel and Kroes 2014).

To put the problem as Coeckelbergh (2015, p. 226) has, ‘by focusing on human–technology relations we might be blind to how technologies such as automation, AI, and robots mediate human–human relations.’ Here, weaponised drones are a relevant example. On paper, the express goal of drones is to operate in situations in which it is considered too dangerous for human pilots to operate. Drones, however, do not just slot into the place of human soldiers, neutrally replacing them, they also allow those situations to become those in which it is too dangerous for human pilots to operate, where an otherwise less dangerous situation might have played out. This is because, with drone technology at their disposal, military operations are now at liberty to produce and enter into conflicts considered too dangerous for human soldiers. This is a case in which the available technology is involved in curating the set of moral decisions available to military generals. Just as the technology is not neutral, nor is its creation and deployment. This is where a consideration of power becomes relevant: if the force employing it only cares about exposing certain lives to increasingly dangerous situations. Concentrated power in the military space thus plays into how technology mediates the moral landscape. Resistance to military intervention is often predicated on the fact that many lives are lost in warfare so wars may become easier to initiate for those who have the power to wage them without the loss of lives they deem valuable (Sparrow 2015).

This similarity in kind, both thought of as ways of describing social relations, urges us to consider the relation between power and moral concern. We have clarified that power has a role in curating moral possibilities and what this means is that it shapes moral relations. What has been outlined thus far has not addressed the extent to which the inclusion of power might aid in shifting us away from moral relativism. If power is as inevitably pervasive as Foucault suggests, then it seems as though by introducing it, we have

only described the complexity of moral relations without getting any closer to grounding moral concern. To make this move, two facets of power are important: power understood as an attempt to control and the disappearance of power (becoming an act of violence) in cases where resistance to power is absent or diminished.

These examples of social relations are clear instances in which the relation is influenced by attempts to control—and generally involve a degree of death and destruction typically understood as morally undesirable, if not entirely impermissible. There is a difference, though, between a relation of power and one of violence. This is a distinction which, when applied to a consideration of moral concern, suggests some basis for which we may deem certain relational features to be morally undesirable/impermissible, thereby actively moving away from the moral relativism implied by Coeckelbergh’s social-relationality.

10 Power and violence: moving away from relativism

Power is not understood here as necessarily nefarious (Maze 2018) and so a consideration thereof cannot on its own address the relativism implied in Coeckelbergh’s social-relationalism. It does, however, allow us to explore what occurs when an attempt to control ceases to be one of power but becomes one in which power is ultimately destroyed, i.e., an act of violence. As discussed aptly by Maze (2018) and Allen (2002), for both Foucault and Arendt, as well as concurring on the aforementioned facets of power, they both understand power to be ‘precisely the promise of potentiality’ (Maze 2018). When the relation between actors/entities ceases to be one in which resistance is possible, this is violence. Violence thus implies a restriction or destruction of potentiality to act otherwise (Arendt 1979, p. 405), ‘in both physical and non-physical terms’ (Maze 2018). According to Foucault, any situation in which the potential to resist is absent, such as the shackled slave, there is no power relation—only an act of violence (Penta 1996; Foucault 1997, p. 292). As described by Oksala (2010, p. 28), ‘[v]iolence can destroy power and politics, but it can never produce them.’ By considering the pivotal role of power in social-relationality, we can also comprehend the potential for and ways in which such relations may constitute violence.

This formulation is a response to the conventional account which holds that domination and violence are fundamental to power (Breen 2007). In contrast, we conceive of violence as the terminus of power (as an attempt to control) and also marks its termination. That is, the capacity to resist power is necessary to its relational materialisation, hence the view that these relations are not necessarily nefarious but allow for the possibility of freedom to resist. By arguing

that violence is not necessary to power, we evade a justification of violence which essentially posits its inevitability. In light of understanding the possibility for relations to be characterised in terms of the extent to which potentiality is restricted, we have some basis to evaluate these relations as either violent or non-violent and thus either permissible or impermissible. While this consideration may not conclusively provide a basis for moral concern, it implies that social-relationality need not entail moral relativism if it takes seriously the degree of capacity and potential (to resist) available in a given relation.

This understanding of power in social relations is not one which evaluates isolated instances to judge moral concern in other instances. That is, there need not be a direct violence occurring in a particular relation (between moral patient and moral agent) in order that it may be deemed impermissible. Given our emphasis on the curation of moral situations, it follows that moral judgements also turn on the relevant historical social relations which have given rise to this particular instance. Specifically, it calls for evaluations of how such circumstances came to be shaped by historical destructions of the potential to act otherwise. This is a natural extension of the point made earlier that Coeckelbergh's social relationality allows for a temporally stretched consideration. That is, we can look to past and prevailing relations to establish whether current and/or future ones may constitute acts of violence. We can apply this line of thought to actions related to the potentiality of future persons when developing technology, where a moral analysis is based on current and past relations of power or violence rather than just the relation between subject and technology. As an example, actions which are so environmentally detrimental such that they may restrict the potentiality of future persons is ultimately an act of violence and may be condemned as such.

We can consider our discussion of power and violence in social relationality as applied to Danaher's (2017a; b) analysis of moral concern in the instance of what he calls robotic rape. Robotic rape is described as sexual (penetrative) engagement with an anthropomorphic sex robot who 'suggest[s] a signal of non-consent' (Danaher 2017a, b). A realist grounding of moral concern, such as Torrance's (2014), would make an argument based on the potential harm which may be inflicted on those deemed deserving of moral concern—in his case—those with consciousness. Although technologically advanced, we do not consider currently existing sex robots as being conscious. Thus, in the case of robotic rape considered by Danaher, the human who is engaging with the robot is the only entity deserving of moral concern and so if they do not incur any harm, on the realist view, robotic rape cannot be deemed morally impermissible. In our variation of the social-relational view, relations are evaluated as of moral concern if they appear to be those of power and they are deemed impermissible if relations

involve violence, arise out of violent relations or justify possibility of future violence. So, engaging in robotic rape can be morally problematised. This is because this robotic relation arises from that of conventional (human) rape which itself is morally reprehensible due to its destruction of potential to act otherwise.³ That is, it is an act of violence. The anthropomorphism of the robot at hand is relevant insofar as it clearly indicates the extent to which human–robot sexual engagement is modelled on human–human sexual engagement, as well as the potential ways it may come to bear on future instances of human–human sexual engagement. The anthropomorphism of the robot in this case is a sufficient, but not necessary, condition to warrant an analysis of power and violence. The fact that robotic rape is modelled on instances of rape and thereby the extent to which it comes to shape our future relations and thus social-relational ethical norms, is of greater significance. Any artifact, and the way(s) it is used may come under consideration if it has some grounding in or bearing on power relations, especially so if there is a risk of perpetuating violent or potentially violent relations.

11 Conclusion and implications

The role of artificial intelligence or technology more broadly in our moral landscape depends upon how this landscape is conceived. The realist theory posited by Torrance which seeks to defend the view that moral concern is grounded objectively comes up short in its capacity to function as an explanatory framework which sufficiently accounts for changing moral sensibilities. On the other hand, Coeckelbergh offers a social-relational theory which, in contrast, argues that moral concern should not rest on the properties of individual entities but on the relations between them. While this view better allows for the consideration of social–historical information about relations, it seems to imply a sort of moral relativism and its focus on how things appear makes it blind to the reality of relations. Crucially, Coeckelbergh's account cannot make sense of the role of power to the extent that it plays out in social relations and curates moral possibilities.

By drawing on an Arendtian and Foucauldian notion of power as an attempt to control a situation and assessing the ways it may function in relation to moral situations, we understand how its presence makes relations morally interesting. Not only this, but a view of power also allows us to identify certain social-relational dynamics as violent.

³ One could also argue, along with Cindy Friedman, that the kind of harm is one in which *we* as agents become the moral patients of our actions towards robots (2020).

We have described violence as a restriction of potentiality, marking the end of a power relation. As we have discussed in relation to technology, this characterisation of social-relational dynamics gives us some basis to say of certain actions or relations that they are morally permissible or impermissible. This assessment retains Coeckelbergh's emphasis on analysing social–historical relations, while allowing for some degree of moral judgement to be made.

While our discussion in this paper responds to work in the philosophy of technology, the outcomes with regard to moral relations and power are relevant beyond this disciplinary discourse. There are two lines of consideration which are beyond the scope of this paper but emerge from the points made. The first relates to the shift away from appearance as conclusively indicative of moral relations. By acknowledging the possibility of moral relations or circumstances whose identification requires more from us than simply “looking out there,” we also understand the explanatory limits of individual phenomenological experiences of power. This calls for careful analysis of moral and power relations, and while the experience of moral relations is important in understanding them, our discussion opens up the possibility that it may be analysed from outside them in a way which is productive and helpful.

A second implication goes along with the previous point and has to do with the idea that power and politics are absent when violence occurs. Following from the problem of going by the way things appear, we might perceive some relation as political when in fact it fundamentally denies agency to some person or group, denying them the possibility to resist. The introduction of this notion of violence to the analysis of moral relations allows us to consider moral relations insofar as they constitute destructions of power. For example, the relation between the corporate employer and the minimum wage employee not in terms of a willingness to work for a stipulated wage but rather in terms of the capacity of the employee to resist participation in such economic activities at all. Rather than depending on the notoriously elaborate and evasive concept of freedom and/or exploring precisely what autonomy entails and where its limits lie, we can focus on exploring the availability of real opportunities and capacities to resist coercion. Due to the clear and basic concern of ensuring these real opportunities to resist, this framework can be used not only in evaluating past and prevailing moral relations but also in the moral dimension of decision making for the future.

Acknowledgements The authors wish to thank each other, and Kiasha Naidoo would like to acknowledge the Centre for Humanities Research of the University of the Western Cape for the fellowship award that facilitated the writing of this paper. The opinions expressed and conclusions reached are her own and not necessarily attributable to the CHR.

Funding Open Access funding enabled and organized by Projekt DEAL. Fabio Tollon is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project 254954344/GRK2073 “Integrating Ethics and Epistemology of Scientific Research”, and he is also also a research fellow at the Centre for Artificial Intelligence Research (CAIR).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alfano M et al (2020) Technologically scaffolded atypical cognition: the case of YouTube's recommender system. *Synthese*. <https://doi.org/10.1007/s11229-020-02724-x>
- Allen A (2002) Power, subjectivity, and agency: between Arendt and Foucault. *Int J Philos Stud* 10(2):131–149. <https://doi.org/10.1080/09672550210121432>
- Arendt H (1970) *On violence*. Harcourt Brace Jovanovich, Florida
- Arendt H (1979) *The origins of totalitarianism [1951]*. Harcourt Brace & Company, San Diego. <https://doi.org/10.1017/S003467050007439>
- Arendt H (1998) *The human condition*, 2nd edn. The University of Chicago Press, Chicago
- Breen K (2007) Violence and power: a critique of Hannah Arendt on the “political.” *Philos Soc Crit* 33(3):343–372. <https://doi.org/10.1177/0191453707076143>
- Brown C (2015) Fish intelligence, sentience and ethics. *Anim Cogn* 18(1):1–17. <https://doi.org/10.1007/s10071-014-0761-0>
- Burr C, Cristianini N, Ladyman J (2018) *An analysis of the interaction between intelligent software agents and human users, minds and machines*. Springer, Netherlands, Dordrecht. <https://doi.org/10.1007/s11023-018-9479-0>
- Césaire A, Kelley RDG (2000) *Discourse on colonialism*. Monthly Review Press, New York
- Coeckelbergh M (2010) Robot rights? Towards a social-relational justification of moral consideration. *Ethics Inf Technol* 12(3):209–221. <https://doi.org/10.1007/s10676-010-9235-5>
- Coeckelbergh M (2012) *Growing moral relations*. Palgrave Macmillan, London. <https://doi.org/10.1057/9781137025968>
- Coeckelbergh M (2015) The tragedy of the master: automation, vulnerability, and distance. *Ethics Inf Technol* 17(3):219–229. <https://doi.org/10.1007/s10676-015-9377-6>
- Coeckelbergh M (2011) Trust, power, and information technology. In: Charles E, Ruth H (eds) *The computational turn: past, presents, futures?* IACAP, Aarhus University, pp 254–256
- Danaher J (2017a) Robotic rape and robotic child sexual abuse: should they be criminalised? *Crim Law Philos* 11(1):71–95. <https://doi.org/10.1007/s11572-014-9362-x>
- Danaher J (2017b) The rise of the robots and the crisis of moral patency. *AI Soc* 0(0):1–8. <https://doi.org/10.1007/s00146-017-0773-9>

- Foucault M (1978) *The history of sexuality: Volume 1: An Introduction*. Pantheon Books, New York. <https://doi.org/10.1353/esc.2015.0079>
- Foucault M (1997) What is enlightenment? [1984]. In: Ranibow P (ed) *Ethics: subjectivity and truth (the essential works of Micehl Foucault, 1954–1984)*. The New Press, New York, p 303
- Foucault M (2001) *Power: essential works of Michel Foucault, essential works of Foucault 1954–1984*. Edited by J. D. Faubion. [https://doi.org/10.1016/S1571-5078\(08\)00422-4](https://doi.org/10.1016/S1571-5078(08)00422-4)
- Friedman C (2020) Human–robot moral relations human interactants as moral patients of their own agential moral actions towards robots. In: Gerber A (ed) *Artificial intelligence research. SACAIR 2021. Communications in computer and information science*. Springer, Cham, pp 3–20. https://doi.org/10.1007/978-3-030-66151-9_1
- Gellers JC (2021) *Rights for robots*. Routledge, London
- Gunkel DJ (2012) *The machine question*. MIT Press, London
- Gunkel DJ (2014) A vindication of the rights of machines. *Philos Technol* 27(1):113–132. <https://doi.org/10.1007/s13347-013-0121-z>
- Gunkel DJ (2018) *Robot rights*. MIT Press, Cambridge. <https://doi.org/10.1126/science.323.5916.876a>
- Johnson DG (2006) Computer systems: moral entities but not moral agents. *Ethics Inf Technol* 8:195–204. <https://doi.org/10.1017/CBO9780511978036.012>
- Maze J (2018) Towards an analytic of violence: Foucault, Arendt and power. *Foucault Stud* 1(25):120–145. <https://doi.org/10.22439/fs.v25i2.5577>
- Müller VC (2021) Is it time for robot rights? Moral status in artificial entities. *Ethics Inf Technol*. <https://doi.org/10.1007/s10676-021-09596-w>
- Oksala J (2010) Violence and the biopolitics of modernity. *Foucault Stud* 10:23–43. <https://doi.org/10.22439/fs.v0i10.3122>
- Penta LJ (1996) Hannah Arendt: on power. *J Specul Philos* 10(3):210–229
- Peterson M, Spahn A (2011) Can technological artefacts be moral agents? *Sci Eng Ethics* 17(3):411–424. <https://doi.org/10.1007/s11948-010-9241-3>
- Singer P (1975) *Animal liberation: a new ethics for our treatment of animals*. New York Review of Books, New York
- Sparrow R (2015) Drones, courage, and military culture. In: Lucas G (ed) *Routledge handbook of military ethics*. Routledge, New York. <https://doi.org/10.4324/9780203148433>
- Tollon F (2020) The artificial view: toward a non-anthropocentric account of moral patiency. *Ethics Inf Technol*. <https://doi.org/10.1007/s10676-020-09540-4>
- Tollon F (2021) Designed to seduce: epistemically retrograde ideation and YouTube’s recommender system. *Int J Technoethics* 12(2):60–71. <https://doi.org/10.4018/IJT.2021070105>
- Torrance S (2008) Ethics and consciousness in artificial agents. *AI Soc* 22(4):495–521. <https://doi.org/10.1007/s00146-007-0091-8>
- Torrance S (2014) Artificial consciousness and artificial ethics: between realism and social relationism. *Philos Technol* 27(1):9–29. <https://doi.org/10.1007/s13347-013-0136-5>
- van de Poel I, Kroes P (2014) Can technology embody values? In: Kroes P, Verbeek PP (eds) *The moral status of technical artefacts*. Springer, Netherlands, Dordrecht
- Villa DR (1992) Postmodernism and the public sphere. *Am Polit Sci Rev* 86(3):712–721

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.