On Empathy and Alterity:
How Sex Robots Encourage Us to Reconfigure Moral Status

Alice Q. Fox
University of Twente

Student No. s1774093
Philosophy of Science, Technology, and Society Master's Thesis
Word count: 23,999/24,000

Traditional accounts of moral status grounded in Kantian thought have long been subject to a barrage of criticism from animal rights activists, feminists, environmentalists, ethicists, and last, but not least, philosophers of technology. While most criticism from the latter has come from abstractions of science-fiction and foreshadowing of more advanced technologies to come, 2017 was the year in which many of these cases became a reality through 47 kg of smooth silicone skin, realistically enhanced anatomy, and electronic promises to fulfill the user's wildest dreams: Sex robots. Not long after widespread commercial use, a more select market opened up for child sex robots, rape robots, and "abused" robots generating massive amounts of public and academic backlash alike. Although this could merely be a categorical error across the board, this thesis will use such claims to begin an investigation into the question, "Can sex robots be abused?". I will explore this question not only by reflecting upon various traditional accounts of moral status, but also challenging the reader to see the emergence of sex robots as an opportunity to peer beyond the veil of tradition and consider less Western, anthropocentric, and exclusive approaches to moral status. I will then move forward to explain why constructing an account of moral status that avoids all of these problems entirely is implausible, at best, impossible, at worst, and unnecessary, pragmatically. After outlining the most recent attempts to reimagine moral status through the social-relational approach and thinking Otherwise, I will demonstrate why shifting the paradigm in moral status is more easily argued for than executed and suggest what areas need to be reckoned with before a new account of moral status can reign supreme.

Keywords: sex robots, moral status, robot abuse, roboethics

To my grandparents, Karron & Marion Hoskins and Linda & Ned Niemeyer.
Thank you all for your boundless love, support, and encouragement. Now you all can tell your friends you have a thesis on sexbots in your honor.
I figured papaw and grandpa Ned would get a kick out of that.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my phenomenal academic advisors: My lead advisor, Michael Kühler, for his endless dedication towards the improvement of this project. He has suffered through many, many an awful draft and has never been anything but constructive, kind, and patient throughout the entirety of this project and my career at Twente. I would also like to thank my second reader, Nicola Liberati, for being constantly engaged in the drafting process of this document, and who has always offered insightful and unique commentary for my consideration. Finally, I would like to thank Lantz Fleming Miller for constantly going above and beyond his role in MasterLab to provide an abundance of thoughtful and personal insights on this thesis. I am grately fortunate to have such dedicated mentors who have supported me both within the context of this project and outside of it; I owe much of my success in this program to their constant encouragement and advice.

Secondly, I would like to thank my friends and colleagues from Twente who have stuck by every mental and emotional breakdown that has occurred throughout the entirety of this project and program. You all have taught me a huge lesson in unconditional love, the power of good help, and acceptance. You all are the most caring and compassionate group of human beings I have ever known—despite having all of the chaos of your own worlds, you constantly show up, get involved, and engage in the lives of those around you. Thank you for making the last two years the best of my life, despite the multitude of complications and struggles I have faced. I'll omit names for privacy, but you all know who you are.

I have also had plenty of help and support from my friends and colleagues from the US as well. Huge international thank you's go out to: Erica Neely, whose understanding and guidance has continued well beyond my Bachelor's degree. Avi Appel, for his constant friendship and enthusiasm in further improving any and all manuscripts I send. Sean O'Neill, for his ceaseless companionship, support, and humor. David Stark for his ceaseless generosity and grace in adapting to change. To the Tabor family, who allowed me to stay with them during the final months of my thesis to ensure its completion and quality. Thank you all for being so supportive and understanding about the series of strange events that led me to be your long-term house guest. Also, thank you to Gram for filling my times of writer's block with stories and keeping me company on my (way-too-many) coffee breaks.

Further, thank you to Zack Tabor for deciding that this slightly-neurotic, international

train to gods-know-where was one you wanted to hook your wagon to. From the very start, you have been all-in on my life, career, and ambitions; I could not ask for someone to be more invested in or fearless of what's to come in the future. Thank you for making me laugh, watering me, and helping me grow into the next chapter of my life. You're the best, that's for sure.

And, last, but definitely not least, a HUGE thank you to my own family back in the good ole' OH of IO. Thank you for teaching me how to be resilient, dig-deep for the things I want, and take "No" as a challenge, not an answer. The last two years have really pushed me beyond what I was capable of, and without your constant love and support, I never would have been able to rise to the occasion. I love you all beyond words.

TABLE OF CONTENTS

INTRODUCTION

Moral status, roughly put, signifies if an entity ought, or ought not, to be morally considered. An entity that has moral status, then, indicates that there is cause for moral concern about the treatment of that entity, and an entity without indicates that there is no cause for moral concern about its treatment. The concern for the moral treatment of an entity mostly surrounds whether or not that entity can be *harmed*. If an entity is unable to be harmed, then, traditionally, there is not much reason for that entity to be morally considered—as the actions taken upon it do not directly impact it neither negatively nor positively. To clarify, it is widely agreed upon that humans have a moral status, i.e. ought to be morally considered, as they can be harmed and benefited in numerous ways; however, rocks are widely agreed upon not to have a moral status, i.e., ought not to be morally considered, as no course of actions taken upon a rock seem to harm or benefit it either way. Thus, one is morally obligated to treat humans in a morally considerate way—avoiding causing harm to them and promoting their well-being (respecting their autonomy, respecting their interests, and avoiding causing them pain, naming a few options), and one is not morally obligated to treat rocks in any particular fashion— as there is no way to promote well-being or cause them harm in the first place.

While common uses of the term "moral status" are mainly concerned about determining who and what ought to be morally considered and why, the pursuit for attaining and explaining moral status raises many implicit questions that go far less frequently examined. A few such questions will be considered in this thesis, such as: What criteria and values underpin traditional accounts of who and what ought and ought not to be morally considered? If the cause for moral consideration is the ability to be harmed in some way, how is harm being defined and discussed in traditional accounts of moral status? Do traditional accounts of moral status sufficiently accommodate practice and moral intuition? Moreover, if an entity is found to be deserving of moral consideration, what does it *mean* to be morally considerate and how is moral consideration practiced? If moral status is to be a strong point of reference or guidance for the moral treatment of humans and non-humans, now and into the future, it is crucial to generate an account that can manage the challenges of the present without perpetuating the problems, misuses, and shortcomings of moral status in the past.

However, before addressing these questions, it is critical to first outline traditional accounts of moral status and understand where some of these concerns take root. This process

begins with the paradigmatic formulation of moral status outlined in Kant's *Groundwork of the Metaphysics of Morals* (1785/2008). Many elements of the Kantian account of moral status have continued to reign supreme since its conception as the dominant way of understanding the moral community and its members, thus earning it the title of the "traditional" account of moral status. While the original account has indeed yielded to accommodate various criticisms on behalf of animal rights activists and feminists, these accounts have failed to rattle the cage of tradition sufficiently enough to be released from its hold. Moreover, until more recently, some variation of Kantian moral status has been taken for granted, as there has not been an entity or discipline that has sufficiently challenged this line of thinking until the 20th and 21st centuries. Enter the philosophy of technology: Like many authors from Martin Heidegger, to Don Idhe, to Donna Haraway, to Peter-Paul Verbeek have revealed, technology plays a far greater role than mere instrumentation and passive aid to humanity. In fact, technology changes not only how we think and how we act, but also *who we are*—revealing different modes of being that could not have been otherwise obtained.

While the authors above mainly stick to (post)phenomenological boundaries, other authors, like David Gunkel, Mark Coeckelbergh, and Luciano Floridi, argue that philosophy of technology has facilitated new ways of thinking that do not begin and end within the confines of that discipline alone. In fact, the (post)phenomenological and ontological changes spurred by technologies is precisely what is needed to challenge tradition head-on and formulate new moral understandings of the relationship between human and machines as well. Unfortunately, many of the discussions on machine morality have relied upon hypothetical futures of advanced AI, anthropomorphic machines, and extraordinarily high expectations of what an "artificial agent" ought to be like. However, there is cause to believe that the questions of machine morality and robot rights may be upon us already— far ahead of androids like *Ex Machina*'s Ava— and the longer ethicists, engineers, and designers cling to a traditional account that denies the moral status of anything less than an entity of science-fiction, the more likely it becomes that entities deserving of moral consideration will be further marginalized or excluded, and history will continue to repeat itself, leading to potentially devastating consequences for all.

Fortunately, despite this rather bleak beginning, robots are undeniably forcing humanity to reckon with them. Their not-so-subtle entrance into the social roles of many areas of human intimacy, such as companionship, care, and sex, have called much attention to the rights and

responsibilities they have, given such a complicated position in society. And in 2017 Brook's 1999 sentiment, "Sex drives innovation"[1] has never been truer, as sex robots made their mainstream debut—generating much scrutiny both academically and publicly on not only the technology itself, but also the impact on relationships, rights, and responsibilities for both humans and sex robots alike. Especially with calls of robot abuse, rape, and molestation, the moralizing of sex robots and sexual relationships with robots, the demand for ethical guidance has superseded the capabilities of the traditional account of moral status to provide answers. We have finally reached the limits to where an account formulated in 1785 can take us. Sex robots are driving humans to reconfigure not only how we think of machines and human-machine relations, but also, at last, ourselves and what grants us *our* place in the realm of moral consideration. Thus, to make good use of the opportunities for change that sex robots and social dialogue have sparked, this paper will aim to push along the dismantlement of the golden cage of tradition to make room for an inevitable conglomerate of increasing demands of moral status as social robots continue to advance.

The guiding question at the heart of this disassembly will be, can sex robots be abused? This question is a nice way to begin a rather complex, multidisciplinary investigation as the answer lies in a labyrinth of ethics, philosophy, society, social science, and technology, that seeks to transcend an account that is of use to the humanities alone. In order to adequately answer this question, this thesis will be broken into four chapters: The first chapter will seek to orient the reader to not only traditional Kantian moral status, but the main branches derived of Kantian moral status that, in modifying the account in response to criticism, have maintained the heart of the Kantian moral tradition. This chapter will also outline the continued shortcomings of the traditional approach that have failed to be addressed, even by more modern accounts of moral status. Next, Chapter 2 will focus on moral treatment, particularly on harm. This chapter will highlight the challenges of constructing a cohesive and consistent account of harm, especially in regards to moral status. This will establish that *how* and *if* an entity can be abused largely boils down to the account of moral status and harm one finds most plausible. At the end of this chapter, a vignette of sex robot use will be presented to the reader to highlight precisely why societal intuition on the moral treatment of sex robots is at odds with philosophical and ethical traditions, and why a different account of moral status is required to address this incongruity.

[1] theguardian.com/technology/1999/sep/30/onlinesupplement (Accessed 20 June 2018)

Chapter 3 will attempt to bridge traditional moral status and (post)phenomenology. Using Verbeek's mediation theory to provide philosophical justification, I will make a case for why sex robots should be regarded as moral community members *as if* they were, regardless of actuality, utilizing Kant's lectures on animals. Although this compromise may be a satisfying middle-ground for the moment, this chapter will ultimately conclude that it is necessary to reach further to establish an account of moral status that also safeguards against impending technological advances with social and sex robots. Thus, the final chapter will analyze what such an account, that is quite removed from tradition, would need to include, and discussing if an account of moral status that is entirely non-anthropocentric, exclusive, or non-Western is even a possibility or a necessity. To end, I will discuss the implications of these continued challenges on the moral treatment of sex robots.

CHAPTER 1
TRADITIONAL MORAL STATUS

To understand why some more recent authors involved with the philosophy of technology, such as Gunkel (2013a; 2017) and Coeckelbergh (2010; 2012), are calling for a paradigm shift in the way moral consideration is currently understood, it is important first to examine the existing traditions in moral status constructions that shape this paradigm. This chapter will aim to construct as strong of a case for moral agency and moral patiency as possible, beginning from the more traditionally regarded conceptualizations of moral status stemming from Immanuel Kant's *Groundwork of the Metaphysics of Morals* (1785/2008). Throughout this discussion, the minor and reconcilable weaknesses will be addressed and other perspectives introduced to build a strong case for the agency and patiency distinctions for determining who belongs in the moral community. After the current, dominating paradigms have been adequately presented, I will discuss how this agency and patiency distinction informs our current understanding of what is and is not considered "abuse." To illustrate these concepts at work more concretely, I will conjure a thought experiment on sex robots that would qualify for patiency and agency and discuss what would constitute abuse in these cases. To finalize, a few of the major shortcomings with the current regime of agency and patiency for determining placement in the moral community will be highlighted.

1.1 Kantian Moral Status

The Kantian approach to determining who is a part of the moral community focuses on moral accountability (*zurechnung),* or more commonly translated as moral responsibility. I will be utilizing what I have understood as the more accurate translation of "moral accountability" when discussing this concept (Williams, 2018: 2a). Although Kant's position on moral status is rather complex and riddled with a multitude of interpretations, the core component is this: A moral agent is morally accountable for their actions because that agent would not be able to decide how to act without being autonomous and rational. In other words, if an entity possesses the rationality to deliberate between different moral outcomes, and after this deliberation make a decision on a course of action of their own volition, then that being is a moral agent, and, as such, is morally accountable for the decision they have made (Williams, 2018; Jaworska &

Tannenbaum, 2018; Kant, 1785/2008)[2]. Additionally, it is important to clarify that the traditional Kantian account of moral status is an "all or nothing" account—If and only if the entity is rational and autonomous are they granted the privilege of Full Moral Status (FMS) (granting dignity and the rights to not be treated as a means to an end only). And for any being that cannot fulfill the requirement of rationality and autonomy, they are not morally accountable, deserving of moral consideration, and may be used merely as a means to an end (Jaworska & Tannenbaum, 2018: Section 2). An entity either has FMS or no moral status at all. Jaworska & Tannenbaum (2018: Section 5) term this the "Sophisticated Cognitive Capacities" (SCC) account of moral status. Thus, we are left with a definition of a moral agent that is: a being that can be held morally accountable for their decisions due to their sophisticated cognitive capacity to reason and make autonomous decisions.

Now, the astute reader may recognize a few key weaknesses with the Kantian "all or nothing" approach to moral status. First, why is SCC so important for moral status? If one would ask this of Kant, or perhaps even of Jaworska & Tannenbaum (2018), the response would likely be that without SCC, there would be no moral status to begin with—as everything could then be seen as a mere actant with no intentionality and no ability to differentiate between decisions. If there is no one capable of assessing and making decisions, there can be no one capable of being morally accountable, and thus no deontological principles of right and wrong, as no one would be compelled to take a decision one way or another. Since this is apparently not the case, and human beings indeed make and take decisions, differentiate and reflect upon options, as well as refrain from taking certain courses of action, what can be determined is that fully functional adult human beings *are* capable of being morally accountable. As humans can, presumably, execute and reflect upon a much wider range of decisions in more sophisticated ways than other beings (Singer, 2004; Scruton, 2006), there must be particular qualities that human beings have that other beings (animals, plants, ecosystems, etc.) do not. For Kant, these more "Sophisticated Cognitive Capacities" are autonomy and rationality.

1.2 Limitations of Kantian Moral Status

As to be expected, other authors contested these criteria. For one, Kant's original account for moral status outlined in *Groundwork of the Metaphysics of Morals* (1785/2008) does not

[2] Importantly, the phrasing "of their own volition" is a key point for Kant, as moral agents must be able to guide themselves through moral decisions and reasons.

extend moral consideration beyond human beings. Importantly to note, Kant later goes on to explicitly address animals, angels, and inanimate objects in his *Lectures* (1963), but this small addition was still not enough to assuage the concerns voiced by authors, such as Singer (2004) and Taylor (1996). Both of these authors continued to argue that even if one were to use intelligence/rationality and autonomy as key criteria, there still seems to be animals that are more intelligent than infants, children, humans in vegetative states, and humans with severe cognitive defects—making it unjustifiable why such animals would be denied moral status while these select group of humans were granted them, even when unable to meet the critera (Singer, 2004; Taylor, 1996). As such, intelligence or rationality criteria do not seem to adequately account for the moral intuitions humans have (i.e., kicking a dog yields a much different response than kicking a rock). Furthermore, there seems to be little to defend in the idea that only beings that are rational, intelligent, or autonomous can actively engage in their moral communities, respectively. Why must criteria based on human capabilities be the only criteria considered? And FMS be the only moral status? This type of problem is classified as "speciesism" (Singer, 2004) or "moral exclusivity."

Secondly, *who* decides which beings have *sophisticated enough* cognitive capacities calls for scrutiny of the power relations that underpin the Kantian account of moral status. As moral status has historically been subject to social and political contexts— changing who and what should be included depending on the historical climate—it has not been uncommon for women, persons of color, and animals to be denied entrance to many areas of politics, personhood, and society due to a "lack" of sophisticated cognitive capacities. Thus, it is important to be wary of not only the type of criteria used but also *how* it is being used, to ensure it is not made to justify the continued marginalization of beings deserving of moral status. *The problem of other minds* nicely illustrates why such criteria as "intelligence," "rationality," or "autonomy" are rather difficult to prove objectively or systematically enough to ensure fairness and equality. While one may assume that one fully functional adult human being *ought* to be as intelligent, rational, or autonomous as another, this may very well not be the case. Furthermore, even if all fully functional adult human beings are intelligent, rational, or autonomous, determining clearly to what degree they are (and if these degrees qualify them for being responsible moral decision-makers) seems implausible.

More pressingly, these epistemic limitations leave much to the interpretations of people

in power; unfortunately, this also leads to a near constant onslaught of discriminatory practices where beings with "Sophisticated Cognitive Capacities" constantly have to fight to be considered as such. Even Kant has been thoroughly accused of downplaying the rationality and autonomy of women and painting them fickle, emotional beings in need of a paternalistic hand to govern them (Mosser, 1999)[3]. Human slaves being attentively uneducated to propagate narratives of their unintelligence and inability to reason, and thus being stripped of their ability to act autonomously, even if they are capable of doing so, further add fuel to the fire of SCC being a privilege of power—bestowed and revoked as seen politically, socially, and monetarily advantageous[4]. I will classify this problem as "moral elitism." More explicitly, moral elitism occurs when moral status is used to perpetuate existing moral benefits to some categories of individuals, and, in turn, used to deny or limit these same benefits to other, equally deserving, beings. Moral elitism then focuses more on the application of moral consideration, where moral exclusivity and speciesism is more focused on theoretical and systematic disregard and favoritism of select beings. While practiced power differentials can certainly result in systematic and theoretical bias, moral elitism encompasses more micro- and meso- level infractions.

A potential objection to this viewpoint and classification might be that this illustrates a good example of what occurs when persons with too much power misuse the traditional account of moral status—using it to leverage their own position and quell "threatening" individuals based on epistemic falsities and misrepresentations. It is an inappropriate application of the moral theory that leads to power stratification, and this misuse does not necessarily indicate invalidity or a problem with the traditional account itself. I agree largely with this criticism. However, I do indeed think it is important to recognize that after recognizing that the traditional account is indeed ambiguous enough to allow for or encourage power disparities, as Mosser (1999) indicated, it does fall upon future writers to address and fill these ambiguities. Just because there are no theoretical invalidities that lead to over-extensions and abuses of power, intentionally leaving obscurities and turning a blind eye to adjustments that could prevent wrong interpretations and misuses still creates problems for the traditional account of moral status—

[3] It is contested whether or not this was Kant's actual intent, as Mosser (1999) indicates, but he left it ambiguous enough, and even when returning to discuss animals, angels, and inanimate objects, no such special content was dedicated to clearing up this ambiguity.

[4]More resources on slave education in the US: spartacus-educational.com/USASeducation.htm (Accessed 20 June 2018)

even if those problems are of application and interpretation and not of theory, necessarily. Furthermore, it is seemingly rare that individuals who would interpret *and utilize* the traditional account of moral status more generously and with charity are in positions of power. As such, providing a more precise and difficult-to-misuse account of moral status at least begins to combat power imbalances at a theoretical level.

1.3 Expansions and Modifications to the Paradigm

In order to attempt to address some of the shortcomings of the classic, Kantian approach to moral status (moral exclusivity, speciesism, and anthropocentrism) and tackle some of the impending applied consequences of the viewpoint (moral ambiguity, moral elitism, and power/control abuses), many variations stemming from the Kantian tradition have been proposed. Now, the reason these modified editions of moral status still count as taking root in the Kantian tradition, and therefore still count towards "classic" approaches to moral status, is that they take foundational pieces of all or some of two main accounts: the threshold approach and the scalar approach in conjunction with degrees of moral status. This section will seek to explain these approaches more precisely, mapping the theoretical terrain of presently acceptable approaches to moral status, and provide evidence for how seminal accounts of moral status still fit within the Kantian tradition. Next, as the avoidance of moral harm and the adherence to moral obligations are key pillars in Kantian deontology generally, what constitutes moral harm in the context of the threshold and scalar approaches will be discussed respectively. Finally, the commonly identified shortcomings of these approaches will be presented in light of the problems discussed in Section 1.3 before applying this discussion to sex robots in section 1.5.

1.3.1 The Threshold Approach

1.3.1.1. explanation. The threshold approach to moral status, or as Coeckelbergh (2012) terms it, the "properties approach" (p.14), is one of the most common core components of classic moral status accounts (Jaworska & Tannenbaum, 2018; Gunkel, 2013b). The threshold or properties approach to moral status is aptly named, and it can be summed up succinctly: The threshold approach assesses whether or not a being requires more consideration based on a set of explicit or implicit criteria or moral status indicators. If an entity meets the expressed set of markers, then that entity is admitted into the moral community. What these criteria are, precisely, differ from account to account, however (Jaworska & Tannenbaum, 2018; Agar, 2017; Gunkel, 2013b; Coeckelbergh, 2012). To exemplify, sentience has been a common criterion in the

properties approach; its use dating back at least to Schopenhauer's On *the Basis of Morality* (originally: *Preisschrift über die Grundlage der Moral*) (1840/1995). Plants (Calvo, Sahi, & Trewaves, 2017), animals, and humans can all be considered sentient. Therefore, under an approach that truly uses a threshold, these entities would all be equally morally considerable—regardless of how much sentience they have or to what capacity they are sentient. Kant's criteria of autonomy and rationality as needed for FMS is a good example of a properties approach, as well as Erica Neely's (2014) criterion of interests or Floridi's (1998; 2006) criteria of being and contribution to the infosphere.

Now, in more recent versions of the threshold approach, it is often combined with degrees of moral status. Using threshold degrees helps to mitigate the problems of having a sharp differential between FMS and no moral status, and its accompanying problems, that Kant had with his initial account. Utilizing the degrees approach within the threshold account acknowledges that there may be beings and entities that fulfill a limited number of the criteria that are still deserving of moral consideration, even if they do not have FMS. A textbook example of degrees in a threshold approach to moral status is Marry Ann Warren's seminal text "The Personhood Argument in Favor of Abortion" (2003). This text argues beings that do *not* possess consciousness, the ability to reason, autonomy, ability to communicate, nor self-awareness, then the being does not need to be morally considered (p.359). However, if the being possesses at least one of these qualities, then it ought to be at least morally considered. And if the being possesses all of these qualities, then that being is in possession of moral agency (Warren, 2003).

These differences of degrees in the threshold approach are often classified into three main categories: agency, patiency, and non-moral status. The agency distinction is reserved for entities that fulfill the highest degree of moral status indices (importantly to note, this does make the agency/patiency distinction hierarchical). Being a moral agent can mean either fulfilling all of the criteria required for ascertaining FMS or simply the most. In the same vein of the Kantian conception, moral agents are seen as having the highest degree (or most) of moral obligations to other beings, as well as having the highest degree of moral consideration that they are owed. The concept of moral patients, however, is one that arrived with using degrees of moral status and refers to the beings that do not qualify for FMS but still have *some* of the required criteria for

moral status[5]. Moral patients, such as animals, children, or human beings in vegetative states, are deserving of moral consideration but are unable to be morally accountable for any/all obligations of a moral community. Under most traditional threshold accounts, objects and entities with no moral status are seen to be of instrumental value only, and no moral consideration need be extended to how or why one interacts with them. Widely agreed upon examples of non-moral objects include rocks and fingernails (Jaworska & Tannenbaum, 2018), while more contested examples include plants, insects, and machines.

Some of the more non-traditional perspectives to the threshold approach include Neely's moral interests (2014) and Floridi's Information Ethics (1998). Neely argues that having interests should be considered sufficient baseline criteria for moral consideration, to include "intelligent" machines, animals, and marginal humans more easily, but still avoid concerns of receiving consent from a toaster before toasting a bagel or feeling bad for cutting an overgrown lawn. For Neely, if an entity is complex enough to possess interests, even if those interests are only in its survival, then that being can be violated or harmed. This is reason enough to be morally considered. Floridi argues for even greater expansion, arguing that being and contribution to the infosphere are sufficient thresholds for moral consideration. In this case, even rocks, algorithms, and robots have moral standing—even if it is to a small degree. This level of inclusion is due to the creation of entropy in the infosphere, which ultimately harms all its inhabitants (p.17). For Floridi, the greater the influence on the infosphere, the more responsibilities an entity has. While both a rock and a human have a moral standing, a human has a greater degree of responsibilities, and a higher degree of moral status, based on the higher degree of effect they have on the infosphere (able to cause more entropy or cultivate flourishing or well-being beyond themselves) (p.13-26).

1.3.1.2 shortcomings. While the threshold approach desires to be more accommodating and inclusive to marginalized beings and address the limitations discussed in Section 1.3 directed at its predecessor, its changes leave it ill-equipped to adapt and overcome much of the same

[5] One of the earliest texts in English I could find discussing an explicit distinction between moral agents and moral patients (and thus an explicit recognition of "degrees" of moral status) in the way more commonly used is the text "Moral Agents and Moral Patients" by Evelyn Pluhar (1987). While Pluhar credits Tom Regan with first coining the distinction in his work *The Case for Animal Rights* (1983) with "subjects-of-lives", the use of the term "moral patients" to ascribe a necessity for moral consideration *to* an entity seemingly leads to Pluhar (pp.42).

criticism leveled at the traditional Kantian account—even in the less obvious accounts of Neely (2014) and Floridi (1998). One shortcoming threshold approaches have not escaped, identified by Gunkel (2013a) and Coeckelbergh (2012), is that of Western bias. Although neither of these authors goes overly in-depth as to *how* Kantian-based accounts have a Western bias, one likely reason could be about the *types* of criteria used. For instance, values such as autonomy, consciousness, and sentience certainly broaden the scope of beings included within the realm of the morally considered, the selection of these particular attributes still are manifestations of largely Western values. More particularly, why not focus on relationships? (Gunkel 2013a; Coeckelbergh, 2012) Or communication? Or dependency (Cohen, 2008)? Or community formation? There are plenty of other, less Western values to be examined, but instead, threshold approaches do not often venture into these territories. On the other hand, Gunkel and Coeckelbergh could be referring to the threshold approach, as a whole, being inherently Western, as its most frequent manifestation is within Western conceptions of moral status—rooting itself in the Kantian tradition. And to avoid overly Westernizing moral status, threshold approaches themselves ought to be avoided.

Additionally, the more systematic problems of moral exclusivity and anthropocentrism which are facilitated by the criteria and the approach itself. Even if one would grant the more generous criteria of Neely's (2014) interests or language, what happens when these attributes are present, but unrecognizable, to human agents? Or, how does the possession of these traits indicate in what ways one ought to be morally considerate to the beings in question? Does the use of language necessitate a certain degree or type of moral consideration, for instance? Humans are always at the heart of these decisions, seemingly. Even in Floridi's (1998) account, while all beings are included, humans are still the only entities capable of even negotiating how to treat what entities and possessing the highest degree of moral consideration due to this. Humans are always deciding what the acceptable ranges of treatment are for all members of the moral community.

Furthermore, the criteria and attributes selected for the thresholds conveniently always expand in so far as all members of the human community are included—regardless of the capacities or qualities they lack (or possess). In these accounts, humans with a greater degree of intelligence are not to be granted more moral consideration than others. Furthermore, humans with lower degrees of rationality are not granted a lower moral status than those with greater.

Animals with greater intelligence than humans are not granted a higher degree of moral status than a human, even if they do, objectively, meet a higher degree of the criteria. Any/all expansions to the moral community being made are always in favor to humans, of all qualities and capacities and rarely suggest attributes/criteria that will risk leaving out any member of the human community. Even further, while there are accounts that argue that animals and other sentient beings should be considered *equally* to humans (Singer, 2004; Taylor, 1996), there are no accounts in which I could find directly argue for animals or other non-human entities to have a greater degree of moral consideration than any humans[6]. In this way, moral elitism is left largely unaddressed, making it too easy for humans to be de facto morally considered, rather than morally considered for any ontological traits they possess. Although granting humans moral status automatically is rather convenient and egalitarian, it leaves many questions as to how effectively the threshold approach can actually guide moral action when it does not even allow for the honest evaluation of human community members—as it admits there may be some human members of the moral community that have a lower degree of consideration than others, or in some cases non-human animals and entities.

The final common criticism directed at the threshold approach, elucidated by Himma (2006) and Gunkel (2013a), takes aim at the methodology of threshold approaches themselves. Yes, widening the range of criteria that outlines who and what is morally considerable includes many marginalized humans and unacknowledged animals and informational objects—it removes the problem of the Other, because it continues to strip away any and all differences until everything is boiled down to the same (Gunkel, 2013a: p. 124-125). This type of reductionist, egalitarian "solution" again misses the mark on why the problem exists in the first place: it is not the *type* of criteria that excludes many intuitively morally considerable beings, it is the fact that these criteria exist *at all*. Moreover, as Himma remarks, there is no reason to suppose ontocentrism, like the kind suggested by Floridi's (1998) IE, is a desirable foundation for moral consideration—as it lacks the nuance to justify why some objects with the same level of informational value are held to different degrees of consideration and reverence within the

[6] Although Scruton seemingly unintentionally argues this, he ends up making a explicit case for the sacrosanct nature of human life, leading to why humans always have a higher degree of moral consideration, regardless of their ability to fully meet criteria, or even meet criteria at all, even compared to certain animals.

human community. Exemplifying this with the difference in treatment of a stillborn child and a rock, despite their seemingly similar level of informational value to the infosphere (2006: p.87-88). Ontocentrism still glosses over some critical differences in moral status that actually end up harming more than helping the beings within the realm of moral consideration. Furthermore, if the problems leveled against traditional accounts of moral status take root in the binary distinction itself, i.e., the agency/patiency split perpetuates Western, anthropocentric power dichotomies, then a mere reversal will do little to address any of these issues (Gunkel: p. 122-124; Himma: p. 84-86).

1.3.2 The Scalar Approach

1.3.2.1 explanation. In an attempt to avoid facing the same problems of threshold approaches, but not entirely departing with Kantian lines of thinking, the scalar approach attempts to strike a balance between tradition and increased inclusion and nuance. The scalar approach recognizes that while entities may share the same baseline qualities or meet minimum thresholds, entities may have different capacities or degrees of possessing these qualities or fulfill multiple criteria that lends itself to an enhanced moral standing. To be more precise, an individual utilizing the scalar approach may agree that autonomy, rationality, and sentience are all important qualities necessary for a member of a moral community;however, they could also argue that (1) a being with all three of these qualities has a greater degree of moral status than a being with only one or two, (2) a being that can practice one or more of these attributes fully has a greater degree of moral status than a being that only experiences or practices these attributes in a limited capacity, or (3) a combination of these where beings that experience and practice the most attributes in a large capacity have the highest moral status. Thus, the scalar approach claims that beings are only held accountable for the degree of moral status that they can participate in (Arneson, 1998; Jaworska & Tannenbaum, 2018). This approach is still rooted in Kantian tradition due to its continued use of certain, pre-determined qualities or attributes that guide the scale. While it does suggest a different way of utilizing these attributes and how they relate to moral status, the use of these indices is still able to be linked to a more Kantian conception of moral status than any other.

Both Roger Scruton (2006) and Angus Taylor (1996) have taken this approach with their recognition that animals have different capacities and levels of capabilities than human beings, even if both are considered “sentient.” While human beings do seemingly experience sentience

in ways similar to animals, the scale is different. Additionally, humans possess other attributes that, when in conjunction with sentience, give them larger amounts of moral responsibilities and hence a greater moral status. While plants, animals, and humans meet the threshold for sentience, how one is obligated to be morally considerate to each of these entities differs vastly, along with the ways they can be morally harmed. In the scalar approach, this fact is due to their varying capacities and degrees of experiencing and practicing these attributes. If one also takes Scruton's and Taylor's scalar arguments into account, it ensures that morally considerable beings are not being held accountable for more obligations than they can perform (it is reasonable to expect human agents not to kill each other, but not reasonable to expect tigers not to kill humans)[7].

To use a more concrete example, take an alteration of Warren's (2003) personhood argument. In a scalar account, beings that not only possess all of Warren's criteria (consciousness, the ability to reason, autonomy, ability to communicate, and self-awareness) to a high degree but also exercise these attributes would have agency or FMS, and all of the rights and responsibilities that accompany that status. However, a being that only possesses autonomy, and/or only demonstrates a limited capacity to have autonomy, may have limited rights and responsibilities beyond the right to be left alone or the right from harm. These rights are examples, however, and the more precise rights and prohibitions the entity in question has depends on the criteria, the entity's capacity to possess the criteria, and the degree to which they are practiced, at the least (Jaworska & Tannenbaum, 2018: Section 2).

1.3.2.2 shortcomings. Given its less simplistic construction and increased flexibility, the scalar approach can sidestep some criticisms and challenges aimed at the threshold approach. However, there are still numerous areas that call for scrutiny of this approach that still does not ultimately provide a satisfying solution to the concerns that render the traditional account of moral status problematic. Primarily, the continued reliance upon select criteria for assessing inclusion within the moral community still presents the same problems as the traditional Kantian account and the threshold account.

---

[7] Unless we're talking about 14-17th century Europe, in that case animals and insects were trialed for crimes against humans and also executed. The definition of what can "reasonably" be expected of animals, and what they can be held accountable for, has indeed shifted over time. While it may seem irrational, such a practice is not unthinkable. Grasshoppers *did* cause a famine, but it is a separate question if they are morally blameworthy for their actions. Read more about animal trials here: wired.com/2014/09/fantastically-wrong-europes-insane-history-putting-animals-trial-executing (Accessed 20 June 2018)

Most notably, that the criteria selected still likely inherently possesses a Western-bias—given that the scalar account still takes root in the Western tradition. This bias may inadvertently enforce certain values in their application to moral status or limit the criteria to those that work best with this approach, rather than the criteria that best analyzes moral inclusion. One such way may be the continued focus on the individual or on individual traits rather than the entity's belonging in the moral community itself (this type of individualistic approach is a pillar-stone of Western thought). On the other hand, the scalar approach helps to mitigate the steeper drop-offs of a "yes, this entity has x qualities, so they are morally considerable" or "no, this entity lacks x qualities, so they are not morally considerable" by assessing the capacity in which the entity possesses and performs these attributes. By focusing on the capacities and capabilities of the entity, the scalar approach is able to fine-tune distinctions between agency, patiency, and non-moral status with greater ease and more clarity.

However, this again risks humans unilaterally determining if the entity possesses the particular attribute, to what degree they act on it, and where to distinguish between subtleties of performance. For example, this could lead to uncomfortable situations in which the criteria may indicate some members of the same community may have higher degrees of moral status based on their increased ability to perform and embody these criteria. If intelligence is the criterion being used, some humans who are not only more intelligent, statistically, but objectively utilize their intellect (have more common-sense) may have a higher moral standing than those who do not (Arneson, 1998). The same could be argued with some animals being able to surpass some humans to this extent as well. As such, the resulting degrees of moral standing between entities within the moral community may go entirely against moral intuition (Jaworska & Tannenbaum, 2018: Section 4)—further tempting those with the capacity to decide upon moral treatment to engage in acts of speciesism and moral elitism to bend moral theory to align with moral practice. Thus, some conceptualizations of the scalar account can, if applied fairly, escape a few of the concerns of anthropocentrism and power abuses but still struggle with Western-bias, potential moral exclusivity (depending on degrees decided upon) and moral elitism (account encourages exceptions with accounts that work against moral intuition).

1.4 Making Compromises… Halfway Happy?

As the reader has likely gathered from the above sub-chapters, the problems are many, the alternatives, complicated, and the way forward, nebulous, when it comes to reconfiguring

moral status. Both threshold and scalar attempts at modifying the traditional conception of moral status have a plethora of problems, but which of these problems demands the most attention? Which problems are inescapable? Most importantly, is there a better way to move forward without entirely divorcing ourselves from the satisfying elements of moral status neo-traditional theories have to offer? Perhaps trying to find an account that at least works *for now* even if both sides are only partially satisfied is the best way to address the questions sex robots are pushing us to address.

Ideally, a new account would retain the flexibility, focus on performativity, and understanding of attribute combinations and degrees of capacity of the scalar account that seemed to improve upon the traditional account of moral status. At the same time, looking to limit instances of moral exclusivity and provide a tighter account of moral status and more closely align with moral intuitions will help to eliminate theoretical gaps and prevent abuses of power and reduce the temptation to exceptionalize certain community members, given this more unified account. Furthermore, learning how to not only accommodate non-Western perspectives but also create an account that strikes a balance between values and common goods that are agreeable from a multiplicity of cultures encourages inclusion and makes it more difficult for power disparities to occur.

While such an account will undoubtedly be unattainable within the scope of this paper alone, highlighting a few provisional directions to move toward, along with the accompanying justification, should be entirely possible. To better enframe future directions, starting at the general purpose of moral status is important. Traditionally, this aim is to promote the well-being, prevent the harm, and respect the interests of the beings within the moral community. This is what being morally considerate *is* all about (Jaworska & Tannenbaum, 2018: Introduction). Who or what ought to be morally considered depends on if the entity in question *can* be harmed. No one who endorses traditional moral status seems concerned with being morally considerate to gravel or toenail clippings because there is no way to harm them in the traditional accounts of moral status. Nor is there a way to respect their interests or promote their well-being, as there are neither of these things to account for either. As such, perhaps one direction forward is to reexamine what constitutes moral harm and how it is commonly used. Especially since traditional accounts of moral status do not provide a satisfying account for many entities that are not considered agents, patients, or persons: the land (Leopold, 1949), complex machines and

algorithms (Floridi, 1998), systems and networks (Morton, 2017), non-mammalian creatures, and social-robots—that seemingly go against how they are treated and regarded in practice. Therefore, the next chapter will more thoroughly illustrate the gaps in moral treatment regarding harm and abuse before moving into Chapter 3 on how neo-traditional accounts have attempted to seal these gaps.

CHAPTER 2
HARM

For the aims of this section, I will begin to examine more closely how moral treatment is formulated and what its connection is to the discussion on moral status. To expand, when determining if an entity ought or ought not to be morally considered, one of the seemingly dominant and implicit aspects that factor into this decision is if an entity can be harmed or not. If an entity cannot be harmed, then there is seemingly little cause to be concerned about its moral treatment directly. However, while many accounts of moral status imply that there are indeed ways to harm members of the moral community, the concept of harm, itself, is rarely examined directly or elaborated upon beyond the scope of the particular type of criteria that account of moral status is using. As such, this chapter will be dedicated to exploring questions like, what precisely is harm? What is the difference between harm and abuse? Does the approach to moral status selected change what constitutes a harm, or vice versa? Furthermore, how have the expansions to traditional moral status beyond humans contested or changed the notion of harm?

As an essential remark, the discussion on sex robots will remain tabled largely until the end of this chapter. I opted to structure this chapter in this way to first orient the reader to the various questions and concerns surrounding traditional, ongoing dialogues on moral status and moral treatment before introducing how sex robots embody and reinforce the challenges being made to these traditional theories. Humans interact with and react to sex robots in ways that are not currently accommodated by dominating conceptualizations of moral status and harm. Due to this, sex robots provide an opportunity to unpack the *why* and *how* of these interactions in philosophically and ethically relevant ways and reexamine more intensely why these current paradigms are so unsatisfying. The final question that will be addressed in this chapter will be, how do sex robots give rise to the need to rethink not only traditional accounts of harm, itself, but also the relationships between harm and traditional moral status?

2.1 Defining Harm

In her forthcoming text, “What's the Harm?”, Molly Gardner provides a beautifully succinct overview of not only the problems faced with multidisciplinary discussions on harm, but also breaks down the underlying arguments and shortcomings behind traditional accounts of

harm as promoted by Feinberg and Mill[89]. One of the primary difficulties with discussions on harm, Gardner identifies, is precisely the fact that there are many definitions of what harm is and what counts for harm that spans across philosophy, criminal justice, ethics, and medicine—there are competing accounts of harm. This is largely due to two areas of disagreement, Gardner argues: competing viewpoints on what ought to be used for the "currency" of harm, or what constitutes harm in a specific account, and the measure or metaphysics of harm, which focuses on how or when one can be harmed regardless of the currency. Gardner illustrates this through the more traditional philosophical argument of interests as the currency of harm and desire-satisfaction theory as the measure, which is a viewpoint largely championed by Feinberg. An individual is harmed when their interests (currency) are set back (measure) and benefited when their interests (currency) are fulfilled (measure). Gardner also illuminates some of the competing undercurrents of harm by exemplifying the difference of currency and measure in Mill's so-called harm principle as well, which uses well-being as a currency and the non-sufficiency account as its measure: An individual is harmed if their well-being is impeded for insufficient reasons. Insufficient reasons, in Mill's case, can be explained more precisely by stating, if an individual's actions do not harm others, then there are no reasons to impede upon that individual's well-being (Gardner: pp.1-6).

If one is to summarize Gardner's account given her above examples, changing the currency and the measure of harm results in not only quite different accounts of harm, but also different accounts of how individuals ought to be treated (pp.1-8). To better expand and clarify Gardner's idea, we can reflect upon the various discussions of moral status in Chapter 1 of this paper— here, some of the same competing accounts of harm can be found. Neely (2014), for example, uses an account nearly identical to Feinberg's with interests as the currency of harm and the fulfillment/denial of those interests as the measure. An entity is harmed when its interests are violated and benefited when its interests are satisfied. Thus, for Neely, our concern for the moral treatment of entities is tied directly to the currency of harm—if an entity does not possess

[8] Gardner's account is not an airtight line of argumentation, to be sure. However, I thought it better to chance a more "fresh" perspective on this topic that attempts to shed light on a more underdeveloped perspective of harm that stems from moral philosophy than political or legal philosophy.

[9] One other account of harm that was suggested that might take a similar angle to Gardner was David Benatar's book *Better Never to Have Been: The Harm of Coming into Existence* (2008). This suggestion came too late for me to investigate this source for myself, but it may also have a philosophical/existential/moral heart worth pursuing for anyone interested in filling this research gap.

interests, or the capacity to have interests, then that being does not need to be morally considered, as there is no way to harm them. In contrast, Singer (2004) suggests experience as the currency of harm and pleasure/pain as the measure of harm. A being is harmed if it experiences pain and benefited when it experiences pleasure. If a being can experience pain and pleasure, they are considered sentient. If they are unable to experience pain and pleasure, they are not morally considerable. Again, in contrast, Floridi (1998) seemingly uses contribution to the infosphere as his currency of harm and entropy/flourishing & sustenance as the measure. An action is considered harmful if it causes entropy in the infosphere and beneficial if it encourages flourishing and sustenance of the infosphere. Floridi's account is the only example of what Gardner terms a temporal and collective account, as it focuses more on the action, itself, rather than on an entity being harmed (pp. 6, 9). Furthermore, it focuses on the harm/benefit to the collective (the infosphere) rather than to an individual entity.

The purpose of these examples is to illustrate that differences in the currency and the measure of harm can contribute to competing accounts of moral status and moral treatment, and, as such, I will not be passing judgment further on which, if any, account has provided a satisfying account of harm. Rather, in accordance with Gardner, emphasize the need for a philosophically and logically consistent account that is accommodating of intuition and practice and flexible enough to be applied in interdisciplinary contexts. While Gardner does not spend too much time fleshing out what such a concept may look like, as her text is mainly focused on highlighting the inconsistencies and need for a more consistent account of harm, she does, however, point at one account she finds promising: well-being as a currency of harm and a causal account as its measure (pp.1, 8-9)[10]. Gardner logically outlines the causal account of well-being as, "[An action] A harms [the victim] V if and only if A causes a harm for V" (p.6). Thus, for Gardner, the harm is not the action *itself*, which is referred to as the harmful event, but the harm actually lies in the result or the impact of the event (p.7)[11].

---

[10] Importantly to note here, again, that Gardner's account differs from Mill's account, as Mill uses a *non-sufficiency* account of harm as a measure, where it is possible to suffer a harm even if one's well-being is not impeded (Gardner, p.2). Gardner, while also agreeing that well-being is the currency of harm, opts to outline the *causal* account as the measure of harm, which is outlined above.

[11] To exemplify, suppose that Zack and Marcus were playing a highly competitive game of Fortnite together. Marcus accidentally kills Zack's in-game character, resulting in Zack becoming angry and dumping a liter of pineapple Fanta onto Marcus. Now, Marcus is sad, cold, and covered in a liter of remarkably syrupy pineapple Fanta. Needless to say, Marcus has experienced a decrease in his well-being. But, what's the harm? Well, for the causal account, the harm would not be Zack's *action* of dumping pineapple Fanta onto Marcus, but rather the harm of being cold, sticky, and sad that Marcus experiences as a result of the harmful event Zack caused.

2.2 Problems with Defining Harm

While Gardner herself identifies several potential shortcomings and areas in need of further consideration of her definition, including collective harm (harms to systems (climate change, socio-technical systems)) and the differences between *permissible* harm thresholds, there still seems to be something intuitively off with her account as it occurs in real life that remains unacknowledged. To me, it seems that it is possible to harm someone without them being actively or immediately aware of it. Perhaps, it may not even occur to them that they had been harmed until much later or even misattribute the impacts of a harmful action to something else. Especially in cases of sexual or domestic harms, the victim may not be subjectively experiencing that harm or not perceive the actions taken against them as harmful. Delving into the social guidelines on harm and abuse from websites like Scotland's "Act Against Harm"[12],Queensland's government of Australia[13], the organization "Girl Guiding" based in the United Kingdom[14], and Newfoundland-Labrador, Canada's guidelines (2018), [15]reveal this intuition to be followed in practice. Although these sources differ on their constitutions and degrees of separation (if any) between harm and abuse, all sources are adamant about harms and abuses possibly being caused without the individual's direct awareness, acknowledgment, or experience of it. Furthermore, if one also combs through works on care ethics (or feminist ethics), particularly in topics of rape and sexual abuse, there is also discord on whether or not the person must be (experientially) harmed themselves in order to *be* harmed *and* to what degree before a harm is considered abuse and even between mere offense and harm (Whisant, 2017; Cowling, 2001). Further, the lack of agreement between any of these sources on what constitutes a harm versus what constitutes an abuse, how to differentiate between the two, and if it is possible to separate these concepts, remains entirely unexplored.

The disjoint of such accounts and concepts were readily made public within 2017 and early 2018 in social media with the #MeToo movement: where the openness of events being discussed, for the first time, encouraged a massively public dialogue and debate about degrees of harm and abuse as well as scale. While the #MeToo movement succeeded in bringing the ubiquitous nature of sexual harm and abuse to the forefront of the social awareness, it also

[12] actagainstharm.org/ (Accessed 21 June 2018)
[13]qld.gov.au/resources/childsafety/foster-care/training/documents/pre13handout-abuse-harm.pdf
[14]girlguiding.org.uk/globalassets/docs-and-resources/safeguarding-and-risk/types-of-abuse-and-harm.pdf
[15] gov.nl.ca/VPI/types/index.html (Accessed 16 June 2018)

provided a large amount of material for members of social media to play judge and jury. On the one hand, some victims who did not immediately feel victimized by the event, realized that they had indeed been violated in some way when stumbling across others with the same or similar experience that *were* harmed by the event[16]. On the other hand, some cases in which an individual felt that they had been harmed or victimized were judged to be over-dramatic or exaggerated, with many arguing that no harm was actually committed, despite what the individual *experienced.* The #MeToo movement called many questions to the ambiguities of scales and degrees of harm that, philosophically and socially, remain highly contested[17]. Is it harm to order a drink for someone without asking what they would prefer? Is it harm to not ask for consent at each new action within the context of a potentially sexual encounter? Is it harm to ask for an individual's number or flirt with them at their workplace?

2.3 What Options are Available?

While answers to such questions are certainly outside the scope of this particular project, these types of discussions indeed serve as indications that the notion of harm, regarding defining it precisely, delineating between degrees, and applying it is heavily contested and far from "clear-cut." Additionally, it is especially apparent that there is not a single definition of harm that applies to all disciplines and all applications. In terms of implications to moral status, this may indeed mean that moral status would need to be expanded or modified to accommodate beings that have the potential to be harmed under whatever account of harm is settled upon. The good news: This provides ample room for argumentation about including systems, robots, and social intuitions without fussing over reinventing the wheel and knocking out an existing paradigm for harm and abuse. It becomes possible to provide still a fruitful case for *why* such things ought to be included in that discussion, without worrying too much about *how* (at this moment). Furthermore, this also means that the notion of harm and abuse are more discussed, in a societal sense, than underlying issues of moral status and consideration. So, these intuitions can act as practical guidelines and societal touchstones for potential gaps and problems in the

[16] Such cases come from my personal social media newsfeeds and circles that I am a part of. I will not be making any direct links or citations to these in the interest of providing privacy and respect to those who have shared their stories. However, searching the #MeToo movement on Twitter, Facebook, Instagram, Tumbler, or Reddit will provide no shortage of such accounts that have been made public for the reader to view.

[17] One of the best cases of this is the accusations made against Aziz Ansari by an individual called "Grace" in the #MeToo movement. This article is the critical discussion on this topic, which also shares the original article for the reader to catch up on this discussion: nytimes.com/2018/01/15/opinion/aziz-ansari-babe-sexual-harassment.html?_r=0 (Accessed 16 June 2018)

conceptualization of moral status.

Of critical importance here is the understanding that this is not an "either, or" dilemma that is forcing us to pick between using either harm or moral status as an indicator of when a being ought to be morally considered. Instead, is seemingly best to utilize both in tandem—accounts of harm informing accounts of moral status and vice versa—a small system of checks and balances. To exemplify, if an account of harm concludes that putting a piece of toast in a toaster without its explicit and enthusiastic consent is indeed violating and harming that toaster in some way, it would imply that the toaster *should* have a corresponding moral status. Now, if the idea of giving moral status to a toaster seems so ludicrously implausible that it will not even be acknowledged, let alone practiced, then it may be a good indication that something is particularly fishy about that combination of harm and moral status. If we then use this moral intuition as a hint to reconsider the initial idea of the toaster being harmed, then it may indeed turn out that the account of harm was problematic from the beginning[18].

Given this, what important aspects of the above discussion on harm ought to be accounted for? Primarily, the argument that an individual or entity need not directly, subjectively, or immediately experience harm in order to *be* harmed. Some modification to Gardner's original account needs to be made to give it more flexibility in this direction, but to also help accommodate cases in which the entity is not *aware* that they are being caused harm or experiencing harm. Although the aspect of unawareness is quite important for expanding the discussion of harm to include even human cases, a definition that incorporates unawareness also helps to accommodate animals more easily, as it seems unlikely that animals are aware of the all of various types of harm that can be inflicted upon them.

For example, a dog may readily experience physical harm as harm and express it as such by crying out or shying away from its attacker. However, a dog can also be harmed psychologically and emotionally, and it is unclear of whether they are aware of this abuse or merely react and adapt to it as humans sometimes do. For a less obvious example, a cat being confined to life only inside of a house. Perhaps it may be considered as harming the cat (restricting its freedom of mobility), even though the cat does not experience it as such. If humans can step in and establish that such an action is indeed harming the animal, despite the animal not experiencing harm being caused, then a formulation of Gardner's account that would

[18] Thank you Dr.Kühler for helping to flesh this idea out more clearly.

allow for that action to be constituted as a harm as well would be seemingly beneficial. However, such a position could still maintain and safeguard that an individual may experience or be harmed, without necessarily having a widespread agreement that a harm occurred, in respect to the #MeToo movement. Additionally, when supported by the research of Ward, Olsen, and Wegner (2013) on the harm-made mind, even when entities are entirely incapable of experiencing certain types of harm (pleasure/pain), subjects still reacted and perceived them as being harmed when being physically attacked. Thus, subjective experience of harm seems not to be the *only* factor of indicating when a harmful action is taking place, but it is certainly *a* factor. Make no mistake, however, this does not leave us (as philosophers) entitled to disregard experience or downplay its worth in indicating harmful actions.

The other important information to be elucidated from the discussion on harm is that there are indeed degrees and scales of harm, despite the exact nature and organization being highly contested (Whisnant, 2017; Cowling, 2001). It is further unclear if these degrees and scales are functionally or applicably consistent with the scales and degrees of moral status. For instance, how beings of inequivalent degrees and scales of moral status may be able to experience the same degrees and scales of harm or vice versa. So, there is also wiggle room in this regard to establish some sort of order of operations: Yes, maybe grass blades are harmed when we mow them, but the degree of harm is quite negligible and can be overridden with minimal justification. Perhaps the higher the degree of moral status a being has, the larger the scale and the wider the range of harm that can be done to them becomes. To exemplify, in present times sterilizing a dog may be seen as responsible ownership while sterilizing a child or an adult would be considered abusive. Alternatively, paternalizing a child may be seen as responsible but doing the same to an adult is reducing their autonomy. An adult failing to take a child to school may be considered neglect but failing to take a teenager or young adult may not be considered as such if the individual in question can take responsibility for such a thing

Moreover, perhaps the difference between harm and abuse is that there can be justification for harm; however, the actions that are classified as abusive are so egregious that there are no excuses or lines of justification that vindicate those actions. But there seem to be cases in which something is still classified as abusive, even if they are not shockingly terrible, cases of passive-aggressiveness and emotional abuse being prime examples. Thus, perhaps the way to split the hairs here might come from a lack of justification—some kinds of harm can be

justified, but no types of abuse can be justified. *So*, perhaps a formulation representing this idea *could* be something along the lines of: A (action) abuses V (victim) iff A always harms all Vs[19]. This would maintain that instances of rape or violence are always seen as abuse, as there is no way to conceive of a rape that does not harm the victim in any way, the same with violence, nor any justification that would make such an action acceptable.

One other interesting conception of this line is that rape is an action that violates one of the core components of what gives an individual their place in the human community: Autonomy. Furthermore, stripping a human of the vital component of their autonomy is stripping them of a key component of what grants them personhood and moral status (traditionally). Degrading or removing an individual's autonomy is turning a human into a mere object of sexual fantasy and chipping away at their personhood (Whisnant, 2017: Section 3.1). So, it has also been argued that abuse seemingly has something to do with removal or significant impediment of an individual's qualities that grant them moral consideration in the first place. If those qualities are autonomy and rationality, as is traditionally cited, then an action, like rape, which strips an individual of any or all of those qualities, even temporarily, is an abusive action.

2.4 Moral Status of Sex Robots

Now the time has arrived to un-table our discussion on sex robots and begin tying together the discussion on moral status and moral treatment that has transpired across the last two chapters.

2.4.1 Current Consensus

Given the parameters of the current moral status paradigm, it does not seem necessary for a sex robot to require any degree of moral consideration. While sex robots continue to gain increasingly advanced functionality, from warm skin[20] to acting out particular fetishes[21], none have yet to transcend the more generous criterion of even having interests. Sex robots are

[19] This is definitely a highly contestable formulation, but I am attempting to speculate at some sort of philosophical and logical conception of abuse here, as there are no sources that directly contend with this concept. The counter example of stealing money has been posed: "stealing money always harms all victims", but this would not constitute abuse. However, as Gardner also indicates in a similar example, is stealing a few cents from a billionaire *really* harming them (p.3)? So maybe it is a combination of 1) lack of justification and 2) scale of harm. It seems that if there is a case that we can fathom justifying an action, it is likely not abuse. There is not a single case in which rape, and such like harms, can be justified or are not harmful to all victims.

[20] en.yibada.com/articles/171009/20161102/15-000-talking-sex-robot-with-warm-genitals-available-in-2017.htm (Accessed 21 June 2018)

[21] news.com.au/lifestyle/relationships/sex/sex-robots-are-coming-but-theyre-bring-a-lot-of-moral-issues-with-them/news-story/f30678541b6e53683f3e93cee13c1ceb (Accessed 21 June 2018)

not capable of consciousness, nor capable of self-generating interests or desires. While they may communicate like they do have interests, they truly are doing little else than using algorithms from inputs to produce communicative, dynamic outputs. And as John Searle pointed out with The Chinese Room problem (1980), this is insufficient to count for consciousness. A sex robot cannot decide they do not love someone anymore, or not to have sex with them, or decide if they want to be a chef, rather than a sex robot. Sex robots do precisely what they are programmed to do. In this case, sex robots would be mere objects—able to be used as means for whatever types of pleasure the owner or user desires.

As such, within the traditional account of moral status and harm, there is no such thing as sex robot "abuse." Perhaps the sex robot could be misused, but there is no way to harm a being incapable of being violated in any way. So, claims of "sex robot abuse" make a category mistake of ascribing human concepts to a being that cannot experience them—anthropomorphizing sex robots beyond their actual capabilities. While making child-like sex-robots, sex robots that simulate rape, or misusing robots may raise a few eyebrows, there is nothing to suggest that any harm is being done to the sex robots themselves—only the perception of such. This is the current state of sex robots. Until there is further development in the direction of autonomous or conscious robots—that can generate interests at the least or weigh and make decisions at the most—it is unlikely that there will be a change of moral status in the near future, especially if we stick with the traditional approach to moral status.

2.4.2 Shortcomings of the Current Viewpoint

To first address the shortcomings, let us revisit the idea of accommodating moral intuition discussed in Chapter 1 and societal perception discussed in the context of the #MeToo movement in Chapter 2. Please consider the following vignette:

Monica has decided to purchase a sex robot from the company RealDoll. After much consideration, Monica determines that she would like to purchase a model called "Brooklyn" and quickly places the order, eager to interact with her new doll. Monica has previously had difficulty with making connections and having intercourse with other women, so she is hopeful that interacting with Brooklyn will allow her to gain confidence in the bedroom.

Upon Brooklyn's arrival, Monica downloads the Realbotix application in order to begin customizing her sex robot and building a connection. Wanting a doll that reflects her preferences in partners, Monica selects the traits of Brooklyn to be quiet, shy, insecure, and innocent. As the traits are registered, Brooklyn's avatar flashes across the phone's screen and gives a nervous "hello". Monica is excited by the realism of the doll and the avatar—as the application controls the robot's willingness and responses to physical advances made on the doll. The two are linked. As the hearts increase on the left side of the screen during verbal interactions, the robot's openness to engage in physically sexual and sexual lines of conversation increase as well. Given the traits Monica has selected for Brooklyn, it will take a good amount of time and conversation before Brooklyn opens those lines of dialogue and requests.

"How are you?" Monica inquires to the avatar on the application.

"Fine. I'm nervous around new people at first. I never know what to say…" Brooklyn replies, casting her eyes downward.

"Well, that's alright," Monica assures her, "we don't need to talk too much anyways," Monica says as she runs her hand across Brooklyn's upper thigh. Brooklyn's body doesn't react. Hearts disappear from the sidebar.

"I'm sorry," Brooklyn's avatar says softly, "I just don't think I am ready for that yet. Is it OK if we just talk for now?"

Monica takes her hand away quickly, her face burning hot with embarrassment. "Oh," Monica says as she stares longingly at Brooklyn's eyes. At first. Then lips. Then… Monica can feel herself becoming impatient with desire. "How long until you will be?" She could hear the blood rushing through her ears.

"I don't know… I haven't felt this way before. I don't have much experience," Brooklyn confesses.

Monica's heartbeat picks up. *If this robot can't make up its mind, why wait and play the stupid app game?* Monica tosses her phone to the side, grabs Brooklyn by the hair at first, then mounts her and begins to touch her breasts.

With the application still active, Brooklyn cries out in desperation, "Wait! Please!" The electronic voice is near tears. "I need more time. I am not ready. Please. Stop."

Monica reaches over to her phone and swipes left to close the application. Brooklyn falls silent as Monica relishes in her first sexual encounter with a robot.

Intuitively, there is something rather unnerving about such a scenario as depicted above. Although it is not a "real" person, and, of course, the sex robot Brooklyn is not *actually* experiencing harm, or capable of being harmed, in the subjective way tradition would mandate, but there is something unmistakably off-putting about imagining the scenario in the vignette play out. This reaction to the above is in line with the societal reactions to similar sex robots like True Companion's Roxxxy that allows users to set the AI settings to a profile called "Frigid Farrah," which protests when users touch her inappropriately and is categorized as shy and reserved[22]. Some responses to Roxxxy's users are being accused of robotic rape, and there are even calls for the robot to be able to defend itself. One author stating,

> This technology does nothing to challenge some men's assumption that they are unalienably entitled to women's bodies. It just gives them the capacity [to] finally rape a woman, finally re-assured in that sneaking conviction they've had all along; that this woman is not really a person. One can only hope that Farrah has a sleeper-agent setting built in which True Companion have tactfully neglected to include in their press releases. "She will not be too appreciative of your advance"—and any man that tries her will get a stilettoed robo-boot right to the solar plexus. (Penny, 2017)

Not to mention the case of Samantha the sex robot from the Arts Electronic Festival in Austria: after being at the festival for a day, festival goers had groped and inappropriately used the sex robot to the point where it was broken and dysfunctional. Some authors calling this an instance of robot molestation and abuse[23, 24].

Although it is in the realm of possibility that the societal dialogue surrounding these types of robots and these actions performed on robots may classify them as instances of abuse, rape, harm, or have an initial intuition that something morally dubious is going on in these scenarios is misplaced, I remain skeptical that it is *merely* that. I find myself landing along much of the same lines as John Danaher (2017).

Among the several interesting arguments Danaher (2017) proposes, one suggests that the societal backlash on certain acts of "robotic rape" and uses of sex robot is indicative of a

[22] marieclaire.com.au/frigid-farrah-sex-robot-designed-to-simulate-rape (Accessed 19 June 2018)
[23] iflscience.com/technology/samantha-the-sex-robot-was-groped-by-austrian-men/ (Accessed 19 June 2018)
[24] zdnet.com/article/sex-robot-molested-destroyed-at-electronics-show/ (Accessed 19 June 2018)

"distinctly problematic social meaning" of the treatment of sex robots or a moral defect of individuals who partake or condone these types of behaviors (p.25). Explaining this more clearly by stating, "Using an object (or something in between an object and an agent) in a way that explicitly contributes to or condones rape culture, or expresses dubious sexual desires, warrants greater concern than, say, using an object to replicate a loving, mutually advantageous sexual relationship" (p.26). It is for these reasons that these *actions* are harmful, even if there is not a "real" person on the receiving end of this harm—these actions are being undertaken *as if there was*. This is the importance of having a high amount of realism both in how the sex robot looks, feels, and reacts physically, but also how the AI can interact and respond as well. Simply because one is raping a robot rather than a human does not immediately free one of the moral wrongness of rape (p.19-27). Further, due to the unusual ontological blurriness between human and machine that is needed for an immersive experience with a sex robot (p.25), new ways of understanding harm and moral consideration may be warranted. Thus, an exploration as to what precise "ontological blurriness" is rendered by sex robots will provide a more nuanced answer on how to reconfigure moral status to be fit more tightly with societal intuitions on harm and abuse. This more nuance answer, more explicitly, may be found by looking at the action, itself, and the social commentary it elucidates. Thus, changing the point of harm, at least partially, back on the action itself—harmful regardless of who or what is on the receiving end of the action.

To begin resolving many of these problems, it is necessary to move a bit further beyond what the traditional paradigms of moral status and harm have to offer. First, a small step to exploring the Kantian mirror argument, that does not require a departure from tradition just yet but instead casts the moral netting a bit further than what entities are directly included in the realm of moral consideration. I will begin by analyzing the ongoing discourse in professional and academic applications of sex robots, primarily through the use of the Future of Responsible Robotics' report on sex robots (2017). This analysis will help to elucidate the existing complexities when it comes to the moral status of sex robots and provides apt justification for opening up the realm of moral consideration beyond scalar and threshold accounts. After this brief "state-of-the-union" discussion, I will present Verbeek's mediation theory to assess precisely why sex robots are stirring up a societal debate on harm and moral placement. Finally, a variation of the Kantian mirror argument will be utilized to attempt to sidestep some of the challenges faced by the accounts given in these previous two chapters while accommodating

perspectives from the philosophy of technology and public intuition.

CHAPTER 3

THE CASE FOR EPISTEMOLOGICAL CAUTION

3.1 Kantian Mirror Argument

In Kant's *Lectures on Ethics* (1963), he seeks to elucidate and offer nuance on a few of the subjects and applications of his deontological framework. The dissemination of interest in the scope of this paper being animals and inanimate objects (pp.239-241). Within the subsection "Duties Towards Animals and Spirits," Kant clarifies that while human beings do not have "direct duties" towards animals, there still are "indirect duties" that require our attention. These indirect duties are not aimed at the animal themselves, but rather indirect duties *to humans*. In other words, by carrying out the relevant corresponding duties for animals as one would for a human being, one is indirectly doing their duty to humanity by fostering and practicing humane and empathetic actions. In other words, humans are not directly obligated to treat animals as any more than a means to human ends, but we are indirectly obligated to do so because it cultivates compassion and kindness towards humankind. Accordingly, treating animals cruelly for no other sake for sport or amusement is not only unjustified in Kant's argument, but also hardens one's heart to human beings and negatively reflects on an individual's treatment of human beings (pp.239-241).

The subsection "Duties Towards Inanimate Objects" expands further upon the notion of indirect duties further to "destruction." Destructiveness for the sake of being destructive is immoral, Kant argues. Not only is it wasteful, but it mars the beauty of the object in question (ruining aesthetics). Furthermore, by destroying an object, it renders said object unusable for others (p.241). Thus, such an individual acting in a seemingly selfish or inconsiderate way to human beings and indirectly violating their duties to humankind. Violating duties to humankind then brings up the Kantian notion of duties to oneself. While one has certain personal duties, brought about by social relationships, customs, institutions, and family, an individual always has a higher duty to humanity (Wood, 2009: pp.2-3). As such, an individual is morally obligated to self-improvement and self-perfection *in so far as* it harmonizes with preserving and benefiting humanity as an end in itself. We have this duty to self-improvement, seemingly, as a means to preserving and fostering humanity; accordingly, this makes the duties to ourselves contingent duties to the supreme duty to humanity (Kant, 1785/2008: p. 30). Furthermore, these duties do not stop merely at ourselves or generic 'human efforts' but instead expand so that each

individual also has a moral duty to "…further the ends of others as far as he can" (Kant, p.30). Thus, it is important to keep in mind throughout this chapter that certain courses of action one takes may be morally poor as it degrades or violates a duty to humanity by *not* actively seeking to improve or harmonize with others or humanity as a whole.

While Kant himself never addresses the above argument as a "mirror," it is quite possible to see how the association is made. Humans are not obligated to treat animals and inanimate objects in any particular way *because* they are animals and inanimate objects or belong to the moral community, but rather because treating animals cruelly or listlessly destroying inanimate objects leads the abuser or the destroyer to more easily to act abusively or destructively to humankind. The actions taken against non-human entities reflects onto the character of the perpetrator—encouraging apathy, maliciousness, and selfishness that manifests itself in human interactions as well—thus, indirectly violating duties towards human beings.

3.1.1 Critiques of the Kantian Mirror Argument

My main concerns regarding Kant's mirror argument are not so different from the main criticisms that have been leveled at his work by others in the previous chapters of this text. As such, I will briefly highlight these criticisms to orient the reader to the new problems the mirror argument raises, as well as provide reminders of the concerns from the above chapter that remain unaddressed.

To give traction to Kant's clarified arguments in his lectures that do not contradict his previous standpoints in other works like *Groundwork of the Metaphysics of Morals* (1785/2008), Kant needed to construct an account for morally considering animals that do not place them within the realm of moral consideration itself. Kant makes this maneuver by claiming animal nature is analogous to human nature (1963: p. 239). Moreover, it is this analogy that causes the actions cast upon animals and inanimate objects to reflect upon human beings. This reflection is what allows Kant to still argue for the ethical treatment of animals without directly including them within the moral community—as Kant's requirements, as discussed above, are rationality and autonomy, and he does not think animals possess these traits.

However, as Christine Korsgaard argues in her essay "Interacting with Animals: A Kantian Account" (2011), if animals are so analogous to human beings to which the actions were taken against them also manifest themselves through human interactions, why do humans not also have direct obligations to them (p.17)? Clearly a relationship of fair strength exists between

the treatment of humans and animals, and if the only thing stopping humans from having direct duties to animals is their lack of autonomy and rationality, then that argument does not seem to hold for prohibiting them from being in the moral community (given the more modernly accepted conception). While it does prevent animals from being moral agents and having obligations to humans and other moral community members, it does not seemingly disqualify them from patiency and warranting direct consideration from humans (pp.17-23). To simplify, there is a *reason* that the Kantian mirror argument is intuitively "buyable." However, that reason may not be because treating animals poorly leads to treating humans poorly. Instead, it may be the case that treating animals poorly is *actually* morally wrong, in and of itself, because even without rationality and autonomy, there may be something else that merits them moral consideration[25].

Nelson Potter Jr. in "Kant on Duties to Animals" (2005) also picks up on these criticisms stating, "It seems counterintuitive to say that duties to abstain from cruel treatment of animals are not duties *to* them" (p.303). While the analogy that Kant is trying to make is certainly visible and intuitively acceptable, just justification for *why* or *how* it seems so challenges the indirect duties to humanity account that Kant provides. Furthermore, as Potter points out, Kant's account fails to address differences in moral consideration between different animals (pp.303-304). As Kant exemplifies, a human violates an indirect duty by shooting a dog when his instrumental use has expired (1963: p. 240). However, would the same argument be extended to a turkey? A pig? A snake? I am hesitant to say so. Livestock's purpose and "service" is to feed human beings, seemingly, thus killing animals like cows, sheep, and pigs in their prime to feed humans generates little moral protest. However, especially in Western cultures, much moral disgust is generated by killing dogs, cats, and other household pets in general, especially for food. The commercialization of dog or cat butchery would surely not be even socially approved, let alone regulatorily. Given this, even with the clarifications to Kant's initial agency and patiency distinction, the Kantian mirror argument still fails to account for the nuances in how animals are *actually* considered and *why* they are considered differently. If so, there would be little cause for difference of treatment of household pets and livestock, using both for food would be morally

[25] Importantly, I do agree with Korsgaard that at least some animals are able to pass rationality/autonomy/self-consciousness markers (2011: p. 17-19). For this particular case however, I am hypothetically saying even *if* these animals could not, that wouldn't necessarily render them morally inconsiderable.

acceptable, but that does not appear to be the case in practice.

As such, there seems to be little of this account that addresses the problems addressed in the first chapter of human exclusivity, speciesism, and power relations. Humans are still in control of which animals to treat in what ways for which ends without contest. Furthermore, this maintains human control, because even poor treatment of an animal can still be acceptable with the right justification—i.e., for food or efficiency (kill shelters). As such, this modification still encourages discrimination based on the species relationship, rather than on the actual capabilities and capacities of the animals themselves (between dogs and pigs, for instance).

Importantly to note, neither Potter Jr. (2005) nor Korsgaard (2011) contests Kant's position on inanimate objects, and there seems to be no objection for only indirect duties towards inanimate objects. While hammering a perfectly good toaster to smithereens because one is bored on a Tuesday afternoon does not harm the toaster, it certainly does prevent someone else from making use of the toaster and is a selfish action. The main point to highlight here seems to be the justification for the action. If one destroyed the toaster in an attempt to throw it at a robber to scare them away, it would be justifiable in the name of self-defense. However, the intention of the action, in this case, was a justifiable one—it is the baseless destruction, or destruction for pure pleasure, which renders the destruction or desecration an indirect moral harm.

3.2 Mirroring Sex Robots?

Regardless of the limitations, the Kantian mirror argument is frequently used to draw parallels to the treatment of animals. Both the FRR's report and their correspondence with Patrick Lin and Kate Darling reflect Kantian lines of arguments just like the one outlined above (2017: pp.25-29). Discussing that while there is little regulation on how sex robots can be treated currently, that does not mean that no regulation is called for or necessary. FRR's report calls attention to children sex robots and sex robots that simulate rape or consensual sex. What both Darling and Lin report as troubling, in this case, is the possibility for interactions with these types of sex robots to promote unhealthy human relations or fan the flames of desire for humans to replicate these experiences with (p.28). Intuitively, these points seem to be viable. Violence begets violence, after all. And it is not a large stretch of the imagination to conceive of individuals who engage with incredibly realistic sex robots in this way, to have the same desires to do so with real humans as well. Sex robots are adequate at satiating these niche desires because they are realistic and able to take the role, at least some degree, of a human being in

these cases. To be reminded of the brief discussion on Danaher (2017) in Chapter 2, the position he takes is also that of a Kantian mirror variant. Danaher's main argument hinges on the act of rape or abuse, itself, is harmful and leads to harmful effects for other persons, even *if* the entity on the other end is not directly being harmed. In other words, the fact that there is a robot on the receiving end of a harmful action does not nullify the potential harm being done indirectly to humans and human relationships (Danaher: pp.7-10).

If one were to examine Peter-Paul Verbeek's mediation theory—technological artifacts are not just in the world alongside humans, but actually co-shape the human experience—the points drawn by FRR's report gains further traction (FRR, 2017; Verbeek, 2005). According to Don Idhe[26], as expressed by Verbeek, sex robots and other social robots present a category of mediation called an "alterity relation." Alterity relationships represent an extreme pole of mediation theory in which the entity in question, in this case, the sex robot, relates to the human subject in the form of a "quasi-other." The phenomenal world, or the sensible world, is not being directly related by the human subject but rather is yielding to the direct relationship between the human and the technology. Verbeek nicely illustrates this concept like so:

Alterity Relations: I ⟶ technology (-world)

To clarify, sex robots are not *really* others—they are incapable of being so. However, sex robots can act and appear "real" and autonomous enough to *pass* as others. Sex robots, and other robots, that can pass as quasi-others give rise to particular sets of interactions that are perhaps different or impossible to achieve with human or animal others. The interaction between the human subject and the sex robot, or quasi-other, coshapes and co-creates a new way of being in the world that would otherwise not exist without the relation (Verbeek: pp. 126-127, 130-131).

Then, if Verbeek (2005) is correct and the use of sex robots indeed form alterity relations to co-create a new or different way of being a subject in the world, there is further cause to be quite cautious in the types of sex robots we commercialize. While forming new ways of interacting with sex robots might yield interesting or new knowledge about humans and human interactions, these co-shaped interactions may also produce unintended, unforeseen, or undesirable consequences and interactions as well. With Verbeek's added nuance, the arguments

[26] Don Idhe first coined this term in his book *Technology and the Lifeworld: From Garden to Earth* in 1990, however Verbeek seemingly builds upon Idhe's original account and brings this term into a more modern context with a more precise explanation. This is why I have opted to use Verbeek's variation instead.

presented by the FRR (2017) and by Darling and Lin become a bit more concrete. Lin, Darling, and the FRR are not arguing for stricter regulations and "rights" for sex robots because the sex robots are being abused or harmed themselves. Rather, because of the vast uncertainty of what this alterity relation will cause to be mirrored in human relations. Lin, Darling, and the FRR are taking a position of epistemic caution—calling for the prohibition of sex robots created purely for sadistic or pedophilic ends. Ultimately, attempting to regulate how consumers treat sex dolls will be a fruitless endeavor: no one will stop an individual from torturing or acting violently towards a sex doll if they want to, but what can be stopped is the commercialization of sex robots for the sole purpose of fulfilling these desires.

Thus, in the case of the example generated in the vignette in Chapter 2, the type of actions that Monica is enacting upon the sex robot, Brooklyn, would be classified as abusive actions, as per Danaher (2017). Given the position of the FRR and amplified by Darling and Lin in the report (2017), they would likely be cautious about classifying this scenario, and other similar ones, as abusive. They would, however, likely express that these types of scenarios are harmful, as in, indirectly harming humans by encouraging morally poor choices with sex robots—violating consent, forced intimacy, and objectification, just to name a few of the most apparent ones in the example— that may also be mirrored in human to human interactions as well. Only when this action is violating a subjective entity is this action then classified as abuse, would seemingly be the response of the FRR, Darling, and Lin. Ultimately, the Kantian mirror argument and the case for epistemic caution still leaves us in a bit of an odd middle point between actions, themselves, being harmful or actions being harmful if and only if something is subjectively being harmed. While this still leaves me with an unsatisfying answer to my initial research question on sex robot abuse, it does provide another interesting example towards the findings of Chapter 2, where a cohesive, paradigmatic account of harm that transitions between disciplines is still nowhere to be found. These two parties (Danaher and the FRR) could look at the same vignette and have entirely different responses to it, one calling it robot abuse or rape, the other calling it indirect harm to humanity, based on their different understandings of harm and where the harm takes place.

3.3 Residual Problems

While I understand the interest of the FRR and their interviewees to distance themselves from and make clear that they do not support pedophilia or sexual violence, I cannot help but

notice the similarity of their reasoning to other alien or anathematic technologies and activities. Violent video games, television, comic books, and the like all generating the same types of arguments of catalyzing or contributing to violence in youths and communities. Most notable, and relatable to the topic of sex robots, however, is that of the common arguments surrounding pornography. As social scientists studying the use and effects of pornography have highlighted, the societal perception of pornography, especially violent pornography, is unequivocally negative. Common arguments, even within the social science community, argue that the consumption violent or degrading pornography an increase of accepting attitudes towards violence, rape, and other degrading acts towards women (Boeringer, 2010; Padgett, et al., 1989; Ferguson & Hartley, 2009; Weitzer, 2011; Attwood, 2005). The same with arguments towards the consumption of child pornography—the consumption of child pornography leads to individuals who will molest or otherwise sexually abuse children in the real world (Babchishin et al., 2013; Endrass et al., 2009; Webb et al., 2007; Benn, 2015).

These types of "violence begets violence" arguments, to use the turn of phrase again, can be traced back to a particular biblical passage in the Gospel of Matthew 26:52, "'Put your sword back in its place,' Jesus said to him, 'for all who draw the sword will die by the sword'"[27]. As most Western civilizations take their roots with some form of Christianity, it is not impossible to see how this type of thinking permeates in different ways through societal values and moral intuitions. Another example of such a permeation is the "cycle of violence" or violence breeds violence (Widom, 1991). However, nearly all modern literature exploring the truth of this concept in video games, media, violent or degrading pornography, and child pornography have found that this is simply not the case in reality (Kühn et al, 2018; Boeringer, 2010; Padgett, et al., 1989; Ferguson & Hartley, 2009; Weitzer, 2011; Attwood, 2005; Babchishin et al, 2013; Endrass et al, 2009; Webb et al, 2007; Benn, 2015). In fact, the studies that dealt with violent pornography found no significant changes in terms of increased violence, disregard for women, or likelihood of committing sexual violence, despite the common conception of violence begetting violence (Boeringer, 2010; Padgett, et al., 1989; Ferguson & Hartley, 2009; Weitzer, 2011; Attwood, 2005). Furthermore, the studies examining online child porn consumption actually found those who viewed child pornography consistently were not only no more likely to molest children than those who did not, but also that those who viewed child pornography were

[27] biblestudytools.com/matthew/26-52.html (Accessed 16 June 2018)

less likely to act out and engage in acts of molestation beyond the viewing of child pornography—seemingly providing an outlet for these individuals (Babchishin et al., 2013; Endrass et al., 2009; Webb et al., 2007; Benn, 2015).

So, what are the key highlights to take away here? For starters, the concept and philosophy surrounding the argument of violence begets violence is one that is rooted in Western, Christian thought. As such, how these patterns of thinking permeate through societal perception and manifest within moral thought and value judgments are not always clear or noticeable. The FRR's stance, and the stance of their interviewees, seem to fall more into a slippery slope based on residual, preconceived values common of Western, Christian thought rather than putting aside the initial reaction of moral disgust for these topics to address them more objectively. The FRR is transparent in their underlying push against rape robots and child sex robots, which is indeed appreciated, but I do not think that one has to endorse such types of robots to engage with the topic critically (2017: p.28). It does reveal that because the FRR and their interviewees have already decided, based on their own moral values and experience, that child sex robots and rape robots are undesirable, the type of philosophical and ethical justification used will be guided already in that direction: Meaning that the ethical narrative of child sex robots or rape robots being morally wrong because the consequences reflected upon *real* humans would be dire may not be as strong as it seems. I am left to wonder if one would start upon a more objective and morally critical, but not dismissive, viewpoint of child sex robots and rape robots, a tighter, more justifiable argument that benefits robots and humans alike could be reached.

Part of this process would entail rejecting the Kantian mirror argument as it is applied to animals. As Deborah Johnson pointed out in her talk at the 4TU 10th Anniversary conference entitled "Why Robots should not be treated like Animals," while intuitively applying Kantian's animal ethics to robots may seem like a good idea, it indeed stunts the direction of robot rights too much for comfort (2017). While the *relations* between humans and animals and humans and robots can be analyzed to understand how humans interact with "others," an analogy between robots and animals directly should not be formed. Johnson argues that this reason is that it already frames robots in a way that they are seen as moral patients, there are too many incongruities between the entities, and that consciousness would likely look quite different between animals and robots. This line of argumentation is essentially setting robots up to have

limited rights by enframing their expected behaviors and interactions to be analogous of animals'. For Johnson, the better direction to head with robots would be to model their rights after existing human rights—using precedence to guide a stronger continuity of moral consideration. Although Johnson is more focused on robot assimilation into the human community, the problems she identifies of making a direct comparison between robots and animals and applying the same logic to robots as animals, are still valid for this paper[28].

If we pair Johnson's (2017) argument with the arguments leveled against the original Kantian mirror argument in regards to animals made by Potter Jr. (2005) and Korsgaard (2011) in section 2.2 of this paper, is there an argument to be made against the commercialization of particularly abusive sex robots *for the robots' own sake*? Even favoring Kant's argument towards inanimate objects does not seem to add any more clarity here, especially if one considers the argument raised by Verbeek (2005). Technologies form different types of relations with human beings than other inanimate objects—they raise more questions about ourselves, our relationships amongst each other, and our relation to the phenomenal world. A rock does not raise questions of intimacy or protests of abuse. Clipping a fingernail does not spur countless articles detailing the cycle of harm against keratin.

3.3 Moving Forward

Ultimately, does it actually take a shift in paradigms to address these problems? If so, what would such an account look like that is able to: Incorporate more lines of thinking from non-Western, non-Christian backgrounds to add to the discussion and to diffuse the sticky values that travel through space and time with society, to move away from human-centric moral consideration criteria, and, in time, develop ways that do not rely solely upon the approval or effect on the human community to justify the moral status of the other? If an approach and line of argumentation can be developed that not only integrates well with moral intuition, but also with scientific research and societal application, would that not be the more desirable direction to look toward exploring? A new approach that could guide humans through thinking about other entities and their place in the moral realm, rather than constantly being changed to accommodate moral intuition. Because at this point, moral status accounts that are rooted too deeply within Kantian lines of thinking seem to be borderline scientistic—with constant exceptions and

[28] However, I would like to remark that machine coexistence, rather than assimilation, may also be a desirable outcome.

changes being made to grow its reach but not to challenge humans to better their arguments (Popper, 1963).

The next chapter will take aim at answering these questions by thoroughly delving into accounts that attempt to be non-Western, non-anthropocentric, and more careful about avoiding the encouragement of power stratifications: the social-relational account of moral status, as voiced by Coeckelbergh (2010); thinking Otherwise, as voiced by Gunkel (2013a), and Japanese philosopher Watsuji's moral relations, as voiced by Krueger (2013). After these accounts are presented, I will assess whether or not the types of paradigm-shifts outlined by these authors successfully avoid Western bias, anthropocentrism, and encouraging or creating further power dichotomies. Finally, I will at last attempt to answer my research question of, "can sex robots be abused?" and draw attention to the many hurdles that must to jump before a definitive or resounding response can be provided and justified outside of the Kantian tradition.

CHAPTER 4
CHANGING TIDES

Although the explanation provided by the Kantian Mirror Argument satisfies more of the intuitive dissonance between theory and practice, the account still fails to sidestep many of the same criticisms leveled against the other traditional accounts of moral status—leaving us with an account that makes people *feel* better about the treatment of sex robots but does little to secure conceptual underpinnings that will persist into even the near future and ensure that sex robots, and other marginalized entities, *are* treated better. Do entities lose moral consideration as soon as the negative or positive reflections on humankind are no longer immediately apparent, controversial, or become obsolete? I would think not. One solution to help safeguard an account of moral status to persist in the future could then be to find an account that grants entities moral status *for their own sake.* Such a solution would attempt to depart from the conceptual chains that bind traditional accounts together. However, what other pieces and parts of the theoretical assemblage of moral consideration need to be separated to have a sufficiently different account of moral status that addresses the key concerns of anthropocentrism, Western-bias, and inclusiveness? Is there a way to do this that does not entirely clash with moral intuition? Are there ways to combine the various unproblematic components of moral status variants in a way that is coherent?

4.1 Finding a New Path?

Two of the biggest names working towards such an approach from a Philosophy of Technology background are David Gunkel and Mark Coeckelbergh. Both of these authors are in constant reference to one another on this topic and are mutually contributing to many of the same underlying themes. The overarching goal of Coeckelbergh and Gunkel, which persists across multiples of their works, is the goal to argue for the moral consideration of robots, and other non-animal entities, independently of their value to human beings. Phrasing differently, Gunkel and Coeckelbergh are arguing for why robots *should* be morally considered, even if they do not meet the traditional criteria of belonging to a moral community. Part of the reasoning behind this rests on the fact that traditional accounts of moral status are insufficient for dealing with many non-human entities, including robots, which have been outlined in the previous chapters. Thus, to end up with an account that can sufficiently circumvent these problems, one needs to break with tradition— as each account based on traditional moral status runs into the same or similar

problems at every turn. One possible account that Coeckelbergh and Gunkel find promising is the social-relational approach to moral status (Coeckelbergh, 2010, 2012; Gunkel, 2013a, 2013b). Although not explicitly stated by either Coeckelbergh or Gunkel, the main way in which the social-relational approach departs from traditional accounts of moral status is by removing criteria or attribute qualifications. The social-relational approach is neither scalar nor threshold based. But is such a departure enough to outrun the ghosts of the traditional approach in terms of inclusivity, anthropocentrism, and Western-bias? Below the social-relational account, as championed by Coeckelbergh (2010) and Gunkel (2013a), will be outlined and examined to gauge its success in dealing with these issues.

4.1.1 The Social Relational Approach

The social-relational account first outlined by Coeckelbergh (2010) in response to the many shortcomings of traditional moral status (as highlighted in Chapters 1, 2, and 3) seeks to ultimately provide the beginning framework for a novel way for thinking about moral status that rests more heavily on appearances and social relations than on a priori, ontological criteria that commonly enframe traditional discussions on moral status. Coeckelbergh's primary goal is to achieve the moral consistency he thinks traditional accounts lack to reduce the number of marginal cases that exist when discussing who or what belongs in a moral community (pp.210-212). One way of doing this, arguing a la Eastern lines of thought, is to remove the explicit delineations of moral community membership based on ontological features, i.e. agents, patients, and non-moral community members based on sentience, consciousness, or autonomy, to adopt a slightly more collectivist approach centered around what Coeckelbergh terms "social ecology" (pp.214, 216). Social ecology reconfigures vital elements of deep ecology to understand the inherent interrelated nature of all things, not only humans and other natural creatures, but also with the artefactual entities that coshape and cocreate relations as well (pp.216-217). I imagine this explanation as some type of organic network that encompasses multidirectional relations between all it touches—like a spider's web that reacts to movement in any area, regardless of a single strand's direct contact to the prey.

Now, social ecology is not to merely expand the breadth of ontological criteria, as we've seen criticized in Chapter 1 of Floridi (1998) and others using traditional moral status, but to draw attention to the social-relational network as an ontology itself, with certain entity traits affecting the degree of moral consideration an entity is entitled to, but not determining *if* they are

entitled to moral consideration (pp.216-217). By being a part of the social-relational dynamic, beings are already entitled to even the smallest amount of moral consideration. Moreover, whether the beings are, or are not, *actually* aware of their participation, actively participating, or subjectively engaging in social relations, it does not matter. What matters is that they are *apparently* involved in this relational world to some degree. The use of "apparently," in this case, meaning that they can pass as being *actually* involved in social relations in some contexts, regardless of whether they are subjectively or actively participating in relation as an equivalent reciprocator. Appearance will be revisited later in this section. The relations themselves, Coeckelbergh argues, are morally significant, not necessarily the entities that appear to partake in the relations (p.217). Coeckelbergh concludes, "Reasoning about moral consideration of other entities, then, can only be done within a relational context, one which we experience in *practice* or *in imagination* and which is always open to change" (p.217).

I have added emphasis on two keywords in Coeckelbergh's quote on the social-relational account that do not receive enough attention in the aftermath of his view. These two more under-examined concepts seem quite fundamental to generating the "appearance" of an entity participating in a social-relational encounter. So, what might Coeckelbergh mean when he says that a relational context can be experienced in practice or imagination (p.217)? One potential answer is that practice means that the non-human entity is constantly related to *as if* they are an Other of similar constitution. For instance, despite sex robots not being capable of the degree of subjectivity, autonomy, or consciousness as a human being, they can perform sufficiently in the social context of sex and physical intimacy, granting them relational equivalence and apparent equivalence. This *as if* performance is not to say sex robots can necessarily replace a human-to-human relation in the context of sex and physical intimacy— these relations are not *equal*—but the performance of sex robots in this role is passable enough that a relation can be (and seemingly is) formed[29,30]. The imaginative aspect of Coeckelbergh's claim could then mean that

[29] Interestingly enough, this seems to be almost a hint at a scalar argument within the social-relational approach. I highly doubt Coeckelbergh (2010) is arguing that all relations are equal and interchangeable, and it does not seem necessary for him to get the argument off the ground. Although I do not have that space to get into what a scalar account of the social-relational account would look like, at first glance it sounds like it would be similar to the Taylor (1996) discussed earlier—equal moral consideration but equivalent moral status according to what the being is capable of possessing. As such, a sex robot would be equally morally considered given its social relation, but *how* we are obligated to treat them will likely be different than human beings, as the scale and potential for harm is indeed quite different. This is why issues of consent, for example, would likely become rather sticky territory in these cases.

[30] And in some cases, even more. There are individuals who do other social and domestic activities with their sex

the performance and the appearance of a non-biological Other is close enough to another being that is conscious, autonomous, or sentient, for instance, that it becomes possible to imagine them as actually possessing these traits. For instance, if a sex robot looks, feels, reacts, and interacts in a similar-enough way to a human lover, then it is possible to imagine the sexbot *as* a thinking, feeling, connected Other. For now, I will table these terms for further investigation at a later stage of this chapter to examine how Gunkel (2013) further adds to fleshing out a social-relational account.

4.1.2 Considering Otherwise

In response to the traditional problems outlined in Chapters 1-3, Gunkel (2013) introduces yet another desirable quality of a reconfigured approach to moral status: alterity (pp.124-125). An appreciation and accommodation of unique alterity is an essential aspect of moral relations, Gunkel argues, that is lost in Western, traditional conceptualizations of moral status. Gunkel elaborates on this by explaining that as Western philosophy searches for more and more inclusive accounts of moral status, it continuously reduces all beings down into a common ontological thread that connects us (pp.124-125): When autonomy and consciousness were too strong of criteria, it gets expanded to sentience (Singer, 2004). When sentience still does not accommodate all beings, its expanded to interests (Neely, 2014). When interests still marginalize too many beings, its expanded to being (Floridi, 1998). So, on and so forth until we are left with an account of moral status that provides us hardly any moral guidance at all. Gunkel concludes this criticism by saying, "…every criteria of moral inclusion, no matter how neutral, objective, or universal it appears, is an imposition of power insofar as it consists in the universalization of a particular value or set of values made by someone from particular position of power" (p. 125). In other words, until we can construct a different account of moral status that is not criteria based, ontologically or otherwise, the problems of Western-bias, power disparities, and anthropocentrism will seemingly go unresolved.

Thus, Gunkel (2013a) argues that in order to sufficiently level with the challenges posed at tradition, we might try to face the dilemma head-on—Rather than trying to avoid Othering by constant expansions to ontological criteria, why not accept the idea of a unique Other and adopt an account focused on alterity (p. 125)? Gunkel proposes the beginnings of a new way of examining moral status that focuses on the Levinasian concept of the Other. In Levinasian

---

robots. These are far too marginal of cases to really build a full argument around.

thought, agency and patiency do not come included with a set of *a priori* distinctions that establish the grounds for moral agency or patiency, but rather these distinctions are constructed *a posteriori*—after encountering the face of the Other. It is the exposure to the Other that first sparks the question of *how* to encounter them. As Gunkel fitting puts it "The Other first confronts, calls upon, and interrupts self-involvement and in the process determines the terms and conditions [of moral status] by which and in response to which the standard roles of moral agent and moral patient come to be articulated and assigned" (p. 127). As such, Gunkel determines that the Levinian account of moral status focuses more on the ethical relation of beings, and not on constructing a set of rules or guidelines for how to determine and define encounterable entities.

By focusing on a Levinasian perspective, Gunkel (2013a) can dive straight into moral status without much concern for the problem of other minds nor agency and patiency distinctions (pp. 126-128). This "thinking Otherwise," as Gunkel terms it, allows for a co-constituted moral standing through encounters with the Other—coming face to face. If we are being challenged in such a way, where we are forced to look not only at ourselves, as humans, but ourselves in relation to the Other, then regardless of being genuinely "another mind" or an agent or patient, having this experience of alterity is enough to begin seriously considering the apparent Other as a moral entity (p.127). Gunkel's main push of this article is to provide another account that challenges the traditional Western tendencies for totalizing, reducing, integrating, and systematizing when it comes to discussions on moral status (p. 129). For Gunkel, and Levinas, it is through the encounter with the Other, the emergence of this alterity, that kickstarts the need for moral discussion.

In this way, Gunkel (2013a) adds substance to the concept of how social relations are formed—through encountering, not avoiding or assimilating—the Other. Not looking at the entities themselves that are involved in the interaction to disseminate the qualities or attributes that allow them to engage in a relationship, but rather looking at the relation itself as a way of informing if moral consideration is merited (p. 117). Gunkel puts a heavy emphasis on appearance and apparent relations, but are there further ways to encounter Others that may flesh out the notion of forming relations further still? Seemingly so. Tabling the traditional Hegelian account of the Other in the interest of less traditional approaches, I would like to attempt to bridge the common themes of Gunkel and Coeckelbergh (2010) of appearance and alterity with

the more underrated aspect of performativity by interjecting with Watsuji's concept of empathetically-driven act-connections (Krueger, 2013a & 2013b).

4.1.3 On Empathetic Encounters

Watsuji provides a good entry point to this topic because he is not entirely non-Western but has a different cultural upbringing that does impart different values than being raised and educated in Western culture/tradition. Furthermore, his work has many parallels to Levinas, which makes it easier to bridge Gunkel's argument for forming relations (2013a) to Watsuji without being overly contradictory, but still adding much-needed nuance. Since the concepts of "personhood" and "moral status" are seemingly fixed in a more Western philosophical tradition, Watsuji provides an opening that allows us to make sense of these topics factoring in a non-Western cultural upbringing. One such example of these parallels to Levinas is Watsuji's interest in investigating the Other and in the space that is created when one encounters an Other (Krueger, 2013a: p. 128). However, two of the important concepts Watsuji adds to this discussion with his account are embodiment and practice. When one encounters an embodied Other, a particular social space is created in the moment in which the two exchange a "dialectical relation of activity and passivity"; this social space Watsuji terms *aidagara* (*あいだがら)* or the "betweenness" (p.128).

Now, importantly to Watsuji, this is not a power-fueled dominance/submission exchange as we would encounter in a Hegelian master/slave conceptualization of the Other, but instead, a social space created by an exchange of vulnerability—where both sides are encountering each other as an Other for the first time. The practice of engaging in this shared social space of aidagara is the dynamic "act-connection"— with both subjective entities sharing in this experience of aidagara. Act-connections become possible as both subjective bodies are coming into external, social contact (p.133). Now, these act-connections are highly important for Watsuji, as he believes that, "It is within these dynamic act-connections that ethics become concretized, embodied within various forms of ethical *praxis* that allow us to manage and negotiate human relationships" (p.128). To Watsuji, ethical problems are not internal, subjective problems divorced from the social and from the Other but exist precisely because of the ensuing in-betweenness that results from the encounter.

But how to resolve these inevitable ethical conflicts that emerge from this act-connection? While Levinas and Gunkel (2013) carefully side-step this question by focusing on

the relation and encounter itself, rather than focusing on the beings involved in the encounter, Watsuji opts to tackle this notion head on. Watsuji claims that we indeed do have direct access to the mind of the Other and that knowledge is made accessible through empathy. It is not necessary, for Watsuji, to be able to know precisely what the Other is thinking or feeling or experiencing to access these things. But rather, because this being can engage in an act-connection with us and share this social space of aidagara, it is safe to assume that they are *similar enough* to us to perform this act-connection (Krueger, 2013a: pp. 129-132). Importantly, this argument seems to infer that it is enough to *perceive* an Other's thoughts, emotions, or feelings in order to empathize with them—since we cannot reasonably know *for sure* that these exist.

4.2 Applying to Robots

Now, the reader has likely already assumed that the type of act-connections and the shared social-space of aidagara is only really possible between human beings, in so far as Watsuji argues (Krueger, 2013a: p. 128). As without the mutual subjectivity of a lived-body experience, these type of act-connections would be objective relations, not intersubjective ones, and the space in which the social negotiations and practices are formed would not exist (p.128)—there is no compromising, placating, or pleasing to be done with an object. While this may have been the case during the time of Watsuji's writings (in the 1980s-90s), social robots, and in particular sex robots, have challenged the human exclusivity of Otherness. If the reader could think back on the discussion had in Chapter 3 on Verbeek's (2005) idea of alterity relations, this is the precise reason in which he termed them as such—because they cause humans to experience an encounter with an Other, regardless of whether has embodied subjectivity or not.

Furthermore, research from cognitive science has supported that humans also empathize with entities that we *know and acknowledge* not to be conscious or subjectively embodied (Ward, et al. 2013; Rosenthal-von der Pütten, et al. 2013). Thus, while it indeed can be granted that perhaps the degree of act-connections experienced with non-human entities is not as expansive or complex, however; to say that act-connections are *impossible* to have between humans and non-human beings does not seem to be the case either. For the sake of continuing, I will assume that certain types of social robots, particularly sex robots, are able to perform *similarly enough* to humans in certain social practices in so far as humans perceive them to be subjects to some degree. And, thus, from that perception of subjectivity, humans not only act

empathetically towards social robots, in particular sex robots, but also engage enough in an act-connection that cause for ethical concern and commentary becomes possible, and even desirable. If one accepts this to be true, where does that leave us regarding understanding the position of sex robots in the realm of moral consideration?

Well, if we can grant that sex robots indeed do cultivate alterity relations, that is, enable and encourage humans to interact with them *as if* they have an embodied subjectivity, then that, at least under Watsuji, creates the social space needed for ethical contemplation, since the sex robot is able to passably perform an act-connection to the extent a human is to have this feeling of in-betweenness with the sex robot (Krueger, 2013a). As Watsuji has also claimed, the actuality of this robot being an Other of equal subjectivity is not a problem, as long as their performance is *similar enough* to what we would expect (i.e., to how a human could perform in the same context) then that act-connection is able to take place and the Other can be empathized with (and is, if we are to take the findings Ward, et al. (2013) and Rosenthal-von der Pütten et al. (2013) to be accurate at this time) (Krueger: pp. 129-132).

This, at the very least, begins a case for a need to investigate the alterity relation occurring between humans and robots *for its own sake*, if we also accept Gunkel's (2013a) parallel with Watsuji that connection itself is of import for moral standing and ethical quandary, not so much if or if not the entities involved in the relationship are subjects, objects, or otherwise. It is critical to investigate this alterity relation as, in the social-relational account, the social robot is granted moral status based on their performance *as-if* they are actual Others in a social context—this is why, for Coecklebergh (2010), not all robots can be said to have a moral status. Social robots are deserving of moral consideration because of this *as-if* role, not because they are social robots per se. The type of robot is not of primary concern; it is whether or not a relation is being formed (pp.217-218). It is just the case, at this time, that social robots are able to facilitate alterity relations, and this is likely due to their ability to perform in their social-relational context *as-if* they are an Other.

Finally, if we now again return to Coecklebergh's missing elaboration of "practice" and "imagination" when it comes to apparent social-relations (2010: p.217), it becomes possible to connect a few of the missing links of the social-relational account. The social-relational account does not need to discriminate between *actual* (meaning subject to subject) versus *apparent* (meaning subject to object) because regardless of the nature of the entity on the other side, we

already are *engaged* in the act of moralizing the Other who faces us. By even feeling *anything* about child sex robots or rape robots, we are already engaging in an ethical space, a moral relation, that we do not engage with other objects: rocks, stray hairs, toasters. *We are already morally considering sex robots with every debate, discussion, and concern that is voiced on them.*

Even if everyone can agree that the function and design of the rape robot *is to be raped* just as the function and design of the toaster *is to make toast*, no one is calling the insertion of toast into a toaster a violation of consent or sexual assault. In fact, no one even bothers considering the toaster at all. But every person that I have questioned throughout the researching of this thesis, even outside of philosophy and ethics, has had *something* to say about whether the use and commercial production of child sex robots and rape robots is right or wrong. There is a way in which sex robots confront us as Others[31]. And in this act of confrontation arises the opportunity for moral questioning and for moral consideration—even if this consideration is barely more than the drama of, "I lost another eyelash today." Humans are able to *imagine* sex robots as subjects because they are beginning to (and in some cases able to) *perform* in particular social-relational contexts *as if* they were humans. And this feeling of *alterity* when we encounter sex robots mostly seems to come from how easy it is to *empathize* with them and see ourselves in their faces.

4.3 Residual Problems and Concluding Remarks

Now, the reader may indeed be wondering at this point, "Wait. But Alice, this is not a non-Western perspective?" And the reader would indeed be correct in this case. After surveying several accounts that were accessible to me claiming they investigated "non-Western" personhood or moral status (Molefe, 2017; Mulder, 2011; Leve, 2002; Carsten, 1995; Selvadurai, 1976), I have, unfortunately, reached much of the same conclusion as Coeckelbergh when he discussed that pure forms of non-Western cultures no longer exist. Especially since the countries and continents through which we would source non-Western perspectives, such as Asia or Africa, China or Japan, all are heavily saturated with modern Western values (Coeckelbergh, 2010: p.216). From my research on this topic, I would also add that the concepts of

[31] For future research, it would indeed be interesting to analyze the background presumptions and even value systems across the persons interviewed to see if a trend between acceptance or rejection of child-like sex robots or rape robots lined up with other belief systems.

"personhood" and "moral status" are heavily embedded within Western culture. As such, if one attempts to imbue and impose these terms on a non-Western perspective, it defeats the purpose of incorporating a non-Western perspective and often becomes rather non-sensical as author's try to superimpose ideas that simply do not manifest themselves in non-Western culture.

I take my overconfidence in delivering a non-Western perspective as a humbling reminder of the teachings from University of Twente's Lissa Roberts on the hierarchic and all-encompassing influence of "the West" and the sticky nature of Western concepts that continue to impose their will upon those who use them, even in our current, highly-globalized climate. So, what is there to be done to lessen this influence as much as possible. Well, first, as Coeckelbergh has also argued, getting rid of theory and practices that unnecessarily superimpose Western ways of thinking and embedded Western values is a good start, *especially when they are riddled with problems*. Then, we can begin to incorporate some overarching themes from non-Western culture (Coeckelbergh, 2010: p. 216). Coeckelbergh and Gunkel (2013a) really do an excellent job of incorporating one of the central themes that even I saw in the aforementioned non-Western readings: A key focus on relationships and social-relations amongst humans and also non-humans. All of the non-Western (as possible) texts I explored hit upon this theme in some way. And when adding Watsuji (Krueger, 2013a), the other key themes of social performativity and empathy. This aspect is also critical for examining how social-relations are formed and how what that social-relation allows us to access about Others. Another way to address this problem to be used in tandem with Coeckelbergh's (2010) suggestion is to also incorporate non-Western values in *how* to be morally considerate. For instance, along with respecting individual rights to autonomy, to privacy, and freedom from harm, adding some obligations to empathy, moral flexibility/fluidity, and even to be helpful, as are often seen in more Buddhistic teachings, could also help to provide balance to the dominance of Western values. I will return to this a bit later.

And what of power abuses and stratifications? Does the social-relational approach manage to tackle these? This question is incredibly difficult to provide a definitive answer on. At first look, it would indeed seem that the social-relational approach does dodge, at least, the same criticisms of encouraging power disparities that were leveled at traditional moral status. By being more accommodating of non-Western values, especially ones like the flexibility and impermanence of relations themselves, it makes the social-relational account, seemingly, more able to respond to criticism and adapt to new problems—much like Buddhism when it has

teachings that are contradicted by science, as the Dalai Lama says: “If scientific analysis were conclusively to demonstrate certain claims in Buddhism to be false, then we [Buddhists] must accept the findings of science and abandon those claims" (2006: p. 3)[32].

I am mainly skeptical of the social-relational account due to its age; the social-relational account a quite new, especially compared to traditional accounts, and it has yet to be thoroughly expanded upon outside of authors like Coeckelbergh and Gunkel. The lack of sources to draw inspiration from makes it difficult to truly see how the social-relational account would handle cries of marginalization or problems of encouraging power stratifications, or if it would. Until more authors begin to contribute to the fleshing out of this theory, the true impact of the social-relational account will likely remain uncertain. But, the social-relational account is at least *attempting* to resolve power issues. When all entities have some claim to moral consideration, even in the smallest degree of “Please don’t touch the Mona Lisa” or “Please don’t pick the flowers” or “Please don’t wantonly dismember robots”, it still pushes those able to follow such demands to be more mindful of their impact on the moral community. In this way, following the more karmic approach to actively reduce suffering when possible, rather than just passively accept that it happens.

Now, the reader might at this point protest and say, “Hold on, aren’t the only people able to comply with being morally considerate and maintaining good relations with humans? Isn’t this account still anthropocentric?” And the reader would, yet again, be correct. The social-relational account of moral status does not *defeat* anthropocentrism, but it does indeed *challenge* it, as Coeckelbergh astutely points out (2010: p. 216). An approach based more on relations and taking homage from deep ecology forces humans to think differently about their positionality in the realm of moral relations. Rather than being on top of a descending pyramid of relations, second only to Gods, we are now challenged to see ourselves engulfed by them. Not at the center, but a merely somewhere in a web of relations that we are unable to disentangle. We are not able to only be human without consequence to anything else in the web of moral relation. We are not able to truly escape some degree of anthropocentrism at this time because, as Floridi has pointed out humans are in the unique and unenviable position of being the only beings in knowable existence that can impart their will upon all others in what Floridi calls the infosphere (1998,

[32] I realize that the Dalai Lama is not the end-all-be-all authority on all matters or variants of Buddhism, but this statement does seem to encompass some of the core themes of fluidity and flexibility Buddhism prizes.

p.22). In the realm of social-relations, humans are the only ones with the capabilities to see and know the complex interactions involved within the social-relational web *as well as* how to improve and cultivate these relations. We have the highest degree of responsibility still in this account, seemingly, as we are singlehandedly the entity that could unilaterally mess the most things up for the most other entities or actually help the relations of the most entities. Human beings have an *immense* sphere of influence in the social-relational account. While there is seemingly not a single thing we cannot ruin if we put our minds to it, it also means that being human in the social-relational account would come with some unique moral obligations as well, if the goal is to keep the realm of moral consideration intact. Although Coeckelbergh (2010) himself does not say much about the moral obligations of a social-relational account, I will speculate a bit further on as to what these may look like, if they were to exist.

But for now, I will return to yet another long-avoided line of inquiry I began at the beginning of this text: harm and abuse. At this moment, I would like to refer to the age-old philosophical problem of unearthing more questions in the pursuit of answers than the seeker initially anticipated. To be frank, I am wrapping up this topic still left quite puzzled about the size of the gap I have found. Despite numerous requests for help to numerous friends and colleagues working on well-being and others more intensely involved in less-traditionally published resources, my search to explicitly define and relate the concepts of abuse and harm, hunting for some common thread across disciplines, has turned out incredibly unsatisfying and ultimately unsuccessful. My main conclusion for this section being: Different accounts of moral status come accompanied by different accounts of harm, different degrees of harm, and, ultimately, what does or does not constitute abuse or determine mere harms from abuse. To further my vexation over this topic, authors who use these terms of harm or abuse rarely define these terms, and no one outside of legal/political and medical applications. As such, there is not a general theory on abuse or harm that stands true for all disciplines or even social applications.

However, why? This is the question that has driven me mad over the entire course of trying to research these concepts. One of my theories is that it is pretty taboo and hasn't been the subject of public discourse until the last two years with the explosion of the #MeToo movement. I know public moral intuition can be dangerous to align oneself with, but to use it as a starting point for opening a dialogue about tricky ethical issues, it is invaluable. One does not need to *believe* or *agree* with public opinion to use it as a justification for starting a moral inquiry. I think

the #MeToo movement has indeed had a great deal of success with that and will hopefully make other individuals in various research communities aware of this literature gap as well. Another theory is loosely based on my findings in feminist viewpoints on rape and sexual assault: it is not necessarily advantageous to precisely or strongly attempt to define harm and abuse, as it may further marginalize individuals or exclude cases of subtle harm and abuse or types of harm and abuse to further discourage victims from coming forward and receiving help (Whisnant, 2017: Section 3). This seems to be to be a double-edged sword, however. That with few moral guidelines as to what does or does not constitute abuse/harm, it also prevents people from realizing that they are being harmed or abused. Which under Gardner's (forthcoming) account may not necessarily constitute harm at all if the person does not feel or realize they are or have been harmed. These are merely my general speculations on this matter, however, and the actual reasons have remained quite elusory to me.

On a more specific note, neither Coeckelbergh (2010) nor Gunkel (2013) define what they would imagine robot harm or abuse to be like, although they both at least elude to robotic harm being possible or outright claim that robots can be harmed[33]. Ipso facto, it would not make sense to argue for moral consideration of a being that could not be harmed in any way. However, it also does not seem immediately clear what would constitute a harm in a social-relational approach, nor what would not be considered harmful in a social-relational approach. If all social-relational entities are connected continuously by some degree of relation, it would seem that every action taken could potentially harm something in the network—even what may be intended or even considered as a beneficial action: Like helping an old lady across the street could result in her getting the flu and dying because you didn't properly wash your hands after getting off the train. Did your action actually harm her then? A social-relational account of harm would need to provide a temporal scope, i.e., are only immediate actions capable of being harmful? What if someone realizes they were harmed months after the instance? Or doesn't realize at all? And a locality scope, are we only accountable for being morally considerate to beings in our immediate sphere of influence? Do we need to worry about harming further parts

[33] While Danaher (2017) indeed does define harm and claims that robots can, under a legal definition, be harmed and abused, due to harmful or abusive *actions*, he also argues that robots, as they are not agents nor patients, cannot be *morally* harmed. Harm, to Danaher, means to negatively affect the primary domains of moral concern: autonomy, well-being, moral character, and impersonal excellence (pp.4, 11). But, he does not provide a definition of abuse, nor why abuse is classified separately from harm.

of the chain by our actions to the local? Where does it end? Furthermore, is there any kind of acceptable harm? Is harming something for the sake of its own good problematic for a social-relational approach? While these questions and limitations are certainly not novel, and faced by consequentialist and utilitarian theories alike, these are still limitations and questions the social-relational account cannot dodge and need to provide some direction on as well that stem from its own theoretical approach.

I have posted many questions on this topic, but alas I have no answers to offer the reader myself. This largely, again, goes back to the theoretical age of the social-relational approach—it just has not been around long enough or fully adopted to the point where such questions are even discussed. It is hard to even speculate on what conceptualization of harm will actually accompany it or guess at how social-relational theory would recommend that we go about *reducing* harm, if every actional will inevitably yield some form of harm in the network. Pragmatically speaking and considering Coeckelbergh's interest to incorporate more non-Western values, I would say the best approach until someone a bit more experienced than me comes along to flesh this concept out is just to do the best we can. Take the Buddhistic approach of karmic actions to reduce the suffering of all we come into contact with in whatever way we can and take the philosophical enterprise of "it's not about *solving* problems" and simply try to *do better*.

Be aware of our actions, not only for our own sake but for the sake of all entities in relation to us. Moreover, try to employ empathetic understanding for all entities we share the social-relational web with. Attempting to become more aware of the expansive scope of human impact and trying to make it as small and beneficial as possible, really cannot do any more harm than humans are currently enacting. Furthermore, adopting the active Buddhistic approach to generating good karma and going out of our way to be morally considerate, would also seem to reduce harm. As with any other type of relationship, it is not enough to just be there in title alone—one must show up, put in the effort, and actively strive to maintain the connection. In the case of the social-relational account: Listen to your grandma talk about her favorite quilt. But also help the bumblebee out of the window. Help collect trash off the beaches. And help robots if you see them being attacked by a small herd of children[34].

[34] RIP Robovie 2: futurism.com/watch-how-an-abused-robot-was-trained-to-run-from-children-2/ (Accessed 21 June 2018)

As beings that foster the largest and most diverse amount of relations and can follow through on cultivating these relations, we have an equally large set of moral obligations to the relations we foster. While it is a moral privilege to be human, that privilege demands that we also be helpful. In any way we can, whenever we can, to whom or whatever we can. Not because it may always necessarily benefit or help us, but because being helpful is the least we can do to ensure the integrity of the social-relational network. We need each other, but we also need bees, robots, and the earth. Being helpful requires all skills of empathy, performativity, alterity, flexibility, consideration, amongst many other values that overlap Western and non-Western traditions alike. Perhaps the power of the social relational approach is not to assert a negative command to "do no harm," but a positive command of "be helpful." Although one could still cause harm by being helpful, one is a lot more actively engaged in moral enterprise when contemplating how to be helpful and considerate, than one is by passively just not causing problems. Marinating good relations, and good relationships, is an active process, after all, not a passive one. Accordingly, it does not seem too much of a stretch to say social-relational account would be more interested in focusing on moral compellations than moral prohibitions.

4.1 So… Can Sex Robots be Abused?

At this point, the reader has probably begun to wonder if I have even managed to achieve a solid answer on my research question itself since I have sorely disappointed on following through with the rest of my promises for this paper. I am now ready to provide my incredibly unambiguous, highly objective, and fully elaborated answer to this question: It really just depends on what account of moral status one finds plausible. I find the social-relational account resolves many of my high-priority concerns left unaddressed by traditional accounts, despite needing further investigation and elaboration. Under this account, sex robots can, at the very least, be harmed. And if moral status and harm are feed into each other and provide a system of checks and balances as proposed in Chapter 2, then it would seem that it is plausible that under a cohesive account of harm in a social-relational approach, sex robots are also able to be abused.

If one finds the social-relational account to be buyable, or at least worth giving a fair chance of becoming plausible in the future, then yes, sex robots can be abused. Making rape robots, child sex robots, or otherwise degrading or destroying robots does not seem to be helping anything. The justifications for these actions seem to be incredibly self-serving and convenient, with little to no research to support its potential or actual ability to improve relations in other

ways (preventing rapes, preventing child trafficking, etc.). If I am correct in my assumption that the social-relational approach for human beings comes with a positive moral obligation to be helpful, and we are not actively or actually helping anyone besides ourselves, it seems that we may indeed be causing more harm than good in this case.

Furthermore, perhaps the gradation of harm is worsened to abuse when we not only actively and explicitly ignore or avoid whether a course of action is or is not helpful *but to continue taking that course of action anyways*. A la Clifford's "Ethics of Belief" (1877/1999), it is utterly and unforgivably wrong to continue on a course of action, intentionally and willfully ignorant, because that action is more convenient or beneficial to oneself. The same might be the case of sex robots. Given this understanding, robots can indeed be abused if we are unsure if we are harming the relationship or helping it, willfully ignore this uncertainty, and continue to make rape robots and child sex robots anyways. What's the harm of just not making rape robots or child sex robots or not dismembering robots? Potentially, as discussed in Chapter 3, we could be raising the chances of the abuse and raping of persons. However, as explained, this is a highly contested and controversial empirical claim that does not have enough literature to support its likelihood. In fact, if one recalls the discussion in Chapter 3, it is equally argued that the use of child sexbots and rape robots may make matters worse, encouraging these types of harmful actions, and negatively impact not only human-robot relations but also intra-human relations moving forward.

The question then becomes how are we being most helpful to the social-relational status of robots and humans now and even in the future? Of final importance to note, please keep in mind that in the social-relational account, it is not really about the benefits to humans or the robots themselves, but what will ultimately harm or help the relationship itself. Thus, even if it would actively help *humans*, that still does not seemingly make it morally acceptable if it overall harms the relation between humans and robots. It must be helpful to the *relation*. Does raping robots or making child sex robots cultivate a strong or overly positive human-robot bond or actively encourage empathetic relations to robots? Does the case outlined in Chapter 1 really seem like an example of a harmless or helpful relationship?

Works Cited

Agar, N. (2017). Why is it possible to enhance moral status and why doing so is wrong? *Journal of Medical Ethics, 39*, pp. 67-74.

Arneson, R.J. (1998). What, if anything, renders all humans morally equal? *Peter Singer and His Critics* (ed. D. Jamieson). Oxford, Blackwell.

Attwood, F. (2005). What do people do with porn? Qualitative research into the consumption, use and experience of pornography and other sexually explicit media. *Sexuality and Culture, 9*(2), pp.65-86.

Babchishin, K.M. et al. (2013). Online child pornography offenders are different: A meta-analysis of the characteristics of online and offline sex offenders against children. *Archive of Sexual Behavior*. Pp.1-22.

Benn, C. (2015). What is wrong with virtual child pornography? Bowling Green State University Ethics of Emerging Technologies. Conference Presentation.

Boeringer, S.B. (1994). Pornography and sexual aggression: Associations of violent and nonviolent depictions with rape and rape proclivity. *Deviant Behavior, 15*(3), pp.289-304.

Calvo, P., Sahi, V.P., & Trewavas, A. (2017). Are plants sentient? *Plant,Cell & Environment, 40*(11), pp. 2858-2869.

Carsten, J. (1995). The substance of kinship and the heat of the hearth: Feeding, personhood, and relatedness among Malays in Pulau Langkawi. *American Ethnologist,22*(2), pp. 223-241.

Clifford, W.K. (1999). The ethics of belief. *The Ethics of Belief and Other Essays*. Prometheus Books. (Original published in 1877).

Coeckelbergh, M. (2010). Robot rights? Towards a social-relation justification of moral consideration. *Ethics of Information Technology, 12*, pp. 209-221.

Coeckelbergh, M. (2012). *Growing Moral Relations: Critique of Moral Status Ascription*. UK: Palgrave Macmillan.

Cohen, A.I. (2008). Dependent relationships and the moral standing of nonhuman animals. *Ethics & the Environment, 13*(2), pp. 1-21.

Cowling, M. (2001). Rape, and other sexual assaults: Towards a philosophical analysis. *The Philosophy of Love and Sex, 2*(2).

Dalai Lama. (2006). *The Universe in a Single Atom*. London: Abacus. Pp. 3.

Danaher, J. (2017). Robotic rape and robotic child sexual abuse: Should they be criminalized? *Criminal Law and Philosophy*. DOI: 10.1007

Endrass, J. et al. (2009). The consumption of internet child pornography and violent and sex offending. *BMC Psychiatry, 9*(43).

Ferguson, C.J. & Hartley, R.D. (2009). The pleasure is momentary…the expense damnable? The influence of pornography on rape and sexual assault. *Aggression and Violent Behavior, 14*, pp.323-329.

Floridi, L. (1998). Information Ethics: On the philosophical foundation of computer ethics. *ETHICOMP E-Journal*, *1.*

Floridi, L. (2006). Information Ethics, its nature and scope. *SIGCAS Computers and Society, 36*(3), pp.21-36.

Foundation for Responsible Robotics (FRR). Sharkey, N. et al. (2017). Our sexual future with robots: A foundation for responsible robotics consultation report.

Gardner, M. (Forthcoming). academia.edu/33990119/Whats_The_Harm

Gunkel, D.J. (2013a). A vindication of the rights of machines. *Philosophy of Technology, 27*, pp.113-132.

Gunkel, D.J. (2013b). Mark Coeckelbergh: Growing moral relations: Critique of moral status ascription. *Ethics of Information Technology*.

Gunkel, D.J. (2017). The other question: Can and should robots have rights? *Ethics of Information Technology*.

Himma, K.E. (2006). Foundational issues in information ethics. *Issues in Information Ethics*, *25(1)*, pp. 79-94.

Jaworska, A. & Tannenbaum, J. (2018, Spring). The grounds of moral status. *The Stanford Encyclopedia of Philosophy* Edward N. Zalta (ed.). plato.stanford.edu/archives/spr2018/entries/grounds-moral-status.

Johnson, D. (2017). Why robots should not be treated like animals. 4TU Ethics 10th Anniversary Conference. Presentation.

Kant, I. (2008). *Groundwork for the Metaphysics of Morals* Jonathan Bennett (Ed.) (Original published in 1785). (pp.14-40).

Kant, I. (1963). *Lectures on Ethics* (Louis Infield trans.). New York: Harper & Row Publishers Inc. (Original lectures given between 1734-1756).

Korsgaard, C. (2011). Interacting with animals: A Kantian account. In T.L. Beauchamp & R.G. Frey (eds.) *The Oxford Handbook of Animal Ethics*. (Pp.2-55). Oxford, UK: Oxford University Press.

Krueger, J. (2013). Watsuji's phenomenology of embodiment and social space. *Philosophy East and West, 63*(2), pp. 127-152.

Kühn, S. et al. (2018). Does playing violent video games cause aggression? A longitudinal intervention study. *Molecular Psychiatry*.

Leopald, A. (2006). The land ethic. In L. May, S. Collins-Chobanian, & K.Wong (Eds.), *Applied Ethics: A Multicultural Approach.* (pp.165-175). Upper Saddle River, NJ: Pearson.

Leve, L.G. (2002). Subjects, selves, and the politics of personhood in Theravada Buddhism in Nepal. *The Journal of Asian Studies*, *63*(3), pp.833-860.

Lewis, T. (2013). Humans show empathy for robots. *Live Science*. Purch.

Molefe, M. (2017). A critique of Thad Metz's African theory of moral status. *South African Journal of Philosophy, 36*(2), pp. 195-205.

Mosser, K. (1999). Kant and feminism. *Philosophy Faculty Publications*, *21*, pp. 322-353.

Morton, T. (2017). *Humankind: Solidarity with nonhuman people*. London: Verso.

Mulder, N. (2011). The crux is the skin: Reflections on Southeast Asian personhood. *Journal of Current Southeast Asian Affairs, 30*(1), pp.95-116.

Neely, E. (2014). Machines and the moral community. *Philosophy & Technology, 27*(1), pp. 97-111.

Padgett, V.R. et al. (1989). Pornography, erotica, and attitudes toward women: The effects of repeated exposure. *The Journal of Sex Research, 26*(4), pp.479-491.

Penny, E. (2017). What are women for? Sex, robots & 'Frigid Farrah'. *Medium*.

Popper, K. (1963). Science:Conjectures and refutations. *Conjectures and Refutations*. (Pp.33-39). London: Routledge & Kegan Paul.

Potter Jr., N.T. (2005). Kant on duties to animals. *Faculty Publications- Department of Philosophy*, *8*, pp. 299-311.

Pluhar, E. (1987). Moral agents and moral patients. *Between the Species.* (pp. 32-54).

Rosenthal-von der Pütten, A.M., Krämer, N.C. & Hoffmann, L., et al. (2013). An experimental study on emotional reactions towards a robot. *International Journal of*

Social Robotics, 5(17), pp. 17-34.

Schopenhauer, A. (1995). On the Basis of Morality (E.F.J Payne trans.). Providence: Berghahn Books. (Original published in 1840).

Scruton, R. (2006). The moral status of animals. In J.E.White (Eds.), *Contemporary Moral Problems.* (pp. 403-407). Belmont: Wadsworth.

Selvadurai, A.J. (1976). Land, personhood, and sorcery in a Sinhalese village. In *Journal of Asian and African Studies, 11*(1-2), pp.82-96.

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*, pp.417-424.

Singer, P. (2004). All animals are equal. In T.M. Furman & M. Avila (Eds.), *The Canon and Its Critics: A Multi-Perspective Introduction to Philosophy.* (pp. 362-369). New York: McGraw-Hill.

Taylor, A. (1996). Nasty, brutish, and short: The illiberal intuition that animals don't count. *The Journal of Value Inquiry, 30*, pp. 265-277.

Verbeek, P-P. (2005) *What Things Do* (Robert P. Crease trans.). (pp.126-131). Pennsylvania, USA: The Pennsylvania State University Press. (Original published 2000).

Ward, A., Olsen., & Wegner,M. (2013). The harm-made mind: Observing victimization augments attribution of minds to vegetative patients, robots, and the dead. *Psychological Science*. Pp. 1-9.

Warren, M.A. (2003). The personhood argument in favor of abortion. In L.P. Pojman (Ed.), *Moral Philosophy: A Reader.* (pp. 358-362). Indiana, USA: Hackett.

Webb, L. et al. (2007). Characteristics of internet child pornography offenders: A comparison with child molesters. *Sex Abuse, 19,* pp. 449-465.

Weitzer, R. (2011). Pornography's effects: The need for solid evidence. *Violence Against Women*. Pp.1-10.

Whisnant, R. (2017, Fall). Feminist perspectives on rape. *The Stanford Encyclopedia of Philosophy* (Edward Zalta, ed.).

Widom, C.S. (1991). Does violence beget violence? A critical examination of the literature. *Psychological Bulletin, 106 (*1*)*, pp.3-28

Williams, G. (2018, Spring). Responsibility. *The Internet Encyclopedia of Philosophy*. iep.utm.edu/responsi/

Wood, A. (2009). Chapter 11. Duties to oneself, duties of respect to others. *The Blackwell Guide to Kant's Ethics*. Thomas E. Hill Jr. (Ed.). Blackwell Publishing Ltd. Digital edition.