

Title: AI Ethics in the Law: Towards a Legal Framework for Sexual Autonomous Robots

Candidate Number: 1025610

Word count: 29,962

Abstract

The academic and public debate surrounding the ascription of legal personhood—viz., the capacity to hold legal rights and/or to bear legal responsibilities—to autonomous robots has so far focused on developing an appropriate legal framework to regulate wrongs caused *by* those robots. In this thesis, I shift the focus: I argue that we need to develop an appropriate legal framework to regulate wrongs inflicted *on* those robots.

I focus on sexual autonomous robots (SARs)—viz., humanoid AI-based autonomous robots which enact sexual relationships with humans—because the imminent emergence of this kind of sexual technology introduces a troubling new way to enact sexual violence. Robotic sexual violence (RSV) cases share the following general structure. There is some human agent A, a SAR B and some action φ such that A φ -s B and, if B were human, A's φ -ing B would be a sexual offence. What is the appropriate legal response to RSV?

There appear to be two classes of responses to this question. On the one hand, there are those who argue for the legal permissibility of RSV by appeal to politically liberal principles. On the other hand, there are those who argue for the legal proscription of RSV by appeal to moral conceptions of the good. My contribution to this literature is to articulate and defend an argument for the legal proscription of RSV by appeal to politically liberal principles which all free and equal citizens in a pluralistic society can reasonably accept. In particular, I argue that there is a Rawlsian “public reasons” justification for the legal proscription of RSV as part of the cultivation of those political virtues which are required by reciprocity in the just society. I show that the attribution of legal personhood to SARs, including legal rights against RSV, would be part of an appropriate legal response to RSV. I conclude with some reflections on the wider ramifications of my thesis for the AI ethics literature.

To Chris, for giving me feet;

To Maria, for appearing.

Do you believe in the human heart? I don't mean simply the organ, obviously. I'm speaking in the poetic sense. The human heart. Do you think there is such a thing?

Ishiguro (2021: 242)

Contents

1. The Robotic Sexual Violence Problem	7
§1.1. Introduction.....	7
§1.2. What are Sexual Autonomous Robots?	9
§1.2.1. Physically Embodying AI	10
§1.2.2. Communication and Interaction.....	11
§1.2.3. Autonomous Capabilities	13
§1.3. The Robotic Sexual Violence Problem (RSVP).....	16
2. Sex, Robots and Political Liberalism	19
§2.1. Liberal Principles and RSV	20
§2.1.1. The Harm Principle	20
§2.1.2. Legal Rights and Legal Wrongs.....	22
§2.1.3. Legal Protections	24
§2.1.4. The Right to Privacy	27
§2.2. Was that (morally) good for you too?	28
§2.2.1. Religion and natural law.....	28
§2.2.2. Feminist Voices	31
§2.2.3. A Question of Virtue	34
§2.3. The Liberal Principle of Legitimacy	37
§2.3.1. The Question of Political Liberalism	38
§2.3.2. Political Conceptions of Justice	40
3. Robots and Reciprocity.....	43
§3.1. The Personal and the Political	45
§3.1.1. Justice Behind Closed Doors: A Question of Reciprocity.....	45
§3.1.2. Beyond the Porn Wars: Rape Myths and the Reporting Crisis.....	48
§3.1.3. Autonomous Pornography	52
§3.2. Virtues of the Polity	55
§3.2.1. When is virtue political?.....	55
§3.2.2. Testimonial Justice as a Political Virtue	57
§3.2.3. Conventions Concerning Discrimination	58
§3.3. Coercing Virtue.....	60
§3.4. Concluding Remarks.....	62
4. Robots and Persons	64

§4.1. The Grounds of Legal Personhood.....	64
§4.1.1. Personhood Legalism and Personhood Moralism.....	65
§4.1.2. Legislative Integrity	69
§4.1.3. Personhood Functionalism	72
§4.2. Oh Brave New World: Towards a Legal Framework for SARs	76
§4.2.1. The Legal Personhood of SARs	76
§4.2.2. Robot Rights	79
§4.2.3. Objections and Replies.....	82
§4.3. Concluding Remarks	86
5. Summing Up.....	87
§5.1. Closing Considerations.....	87
§5.2. Conclusion	90

1. The Robotic Sexual Violence Problem

a dissociation of sensibility set in, from which we have never recovered

Eliot (1921)

§1.1. Introduction

In 2017, the European Parliament called upon the EU Commission to table a legislative proposal laying down a set of civil law rules on robotics and artificial intelligence (AI). The most striking recommendation in the Parliament’s final report is that of

“creating a specific legal status for robots [...] so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause [...]” (Delvaux 2017: 18).

The Parliament’s proposal faced substantive criticism and was eventually abandoned.¹ But legal personhood—that is, the (legal capacity for the) holding of legal rights and/or bearing of legal duties²—for autonomous robots has been a live topic of debate in the philosophical literature for several decades.³ In particular, in unison with the Parliament’s recommendation, a recent surge in academic interest in ascribing legal personhood to autonomous robots reflects and responds to the widespread concern that autonomous robots will generate morally and legally problematic

¹ For criticism of the Parliament’s proposal: Nevejans (2016); European Economic and Social Committee (2017); Karner (2019); Wagner (2019); Riehm (2020); Eidenmüller and Wagner (2021). Evidence of the proposal’s abandon appears in the European Parliament Resolution (2020/2014[2021]: ¶7).

² There are different ways to cash out this orthodox view (see, generally, Smith 1928; Kurki 2019). These include the bearing of rights *or* duties (Gray 1997); the bearing of rights *and* duties (Bilchitz 2009: 42; *Nonhuman Rights Project v. Lavery* [2014]: 5); the *capacity* for bearing rights/duties (Wise 2010: 17; *Restatement (Third) of Agency* 2006: §1.04; cf. Rochfeld 2013: 15; Larenz 1983: 41) or for participating in legal relations (von Savigny 1840, 1867, 1884; Wise 1998: 801; Salmond 1920). See also Kelsen’s minimalist approach to legal personhood: Kelsen (1919; 2005; 1945). I assume the orthodox view in this thesis, although it is not without its dissidents (see, e.g., Kurki 2019).

³ The *locus classicus* is Solum’s landmark (1992). But the earliest academic mention of this kind of approach appears in Stone (1972[2010]: fn.26). Legal personhood for AI systems is also discussed in Koops *et al.* (2010), critiqued in Bryson *et al.* (2017) and Santoni de Sio and Mecacci (2021), but advocated for in Turner (2019) and Chesterman (2020). In the media, see, for example: Hern (2017).

“responsibility gaps”, namely cases in which an autonomous robot causes harm or damage for which no human agent can justifiably be held responsible.⁴

In this thesis, I shift the focus of this discussion in the following way. I develop a new argument for the ascription of legal personhood to certain kinds of autonomous robots. But my argument focuses not upon what these robots might do to *humans*, but rather upon what humans might do to *them*. To advocate for the ascription of legal *rights*, rather than legal *responsibilities*, to a certain class of autonomous robots is therefore my ultimate aim in this thesis.⁵ I focus on sexual autonomous robots (SARs)—viz., humanoid AI-based autonomous robots which enact sexual relationships with humans. This is because the recent literature reveals that the imminent emergence of this kind of sexual technology introduces a troubling new way to enact sexual violence, which calls for the development of a legal framework for SARs.⁶ This thesis is my contribution to that broader project.

In this chapter, I foreground the challenge posed by what I call the “robotic sexual violence problem” (RSVP). Robotic sexual violence (RSV) cases share the following general structure. There is some human agent A, a SAR B and some action φ such that A φ -s B and, if B were human, A’s φ -ing B would be a sexual offence. The RSVP is as follows: what is the appropriate legal response to RSV cases?⁷ In chapter two, I explore two classes of responses to the RSVP which have emerged in

⁴ Matthias’ (2004) notion of responsibility gaps has been especially influential in debates surrounding lethal autonomous weapons—viz., “weapons system[s] that, once activated, can select and engage targets without further intervention by a human operator” (US Department of Defense 2012). See, for example: Sparrow (2007, 2016) Asaro (2012, 2016); Crotoft (2015); Horowitz and Scharre (2015); Roff (2013); Galliot (2015). But responsibility gaps are a problem for autonomous systems more generally (Amoroso & Tamburrini 2019; Mittelstadt *et al.* 2016; Santoni de Sio & Mecacci 2021; Kiener 2022).

⁵ Rights and responsibilities are, however, intimately intertwined in law, as I return to in chapter five.

⁶ See, for example, the first consultation report by the Foundation for Responsible Robotics, entitled “Our Sexual Future with Robots”, calling for the urgent development of a legal framework for SARs (Sharkey *et al.* 2017; E&T editorial staff 2017).

⁷ Notable recent contributions specifically concerning the RSVP include: Chatterjee (2020); Danaher (2017a, 2017b); Sparrow (2017); Strikwerda (2017). The RSVP has garnered political attention too: the USA House of Representatives and the UK’s Crown Prosecution Service recently banned the importation of child-like SARs and sex dolls. Note that

the SARs literature. On one hand, there are those who argue *against* the legal proscription of RSV by appeal to politically liberal principles which all free and equal citizens in a pluralistic society can reasonably accept. On the other hand, there are those who argue *for* the legal proscription of RSV by appeal to some moral conception of the good. This illuminates a lacuna in the recent literature: no one has argued *for* the legal proscription of RSV by appeal to politically liberal principles. It is to the remedy of this omission that I turn in chapter three. Here, I articulate and defend the first politically liberal argument for the legal proscription of RSV: I contend that RSV is incompatible with a core value of political liberalism, one which Rawls (1993[2005], 1997) calls “reciprocity”. In chapter four, I suggest that the ascription of legal personhood to SARs, including legal rights to be secure from RSV, would be a fruitful starting point for the implementation of the legal proscription of RSV. I conclude in chapter five with some reflections on the ramifications of my arguments for wider debates in the AI ethics literature. To begin with, though, I wish to clarify the parameters within which my arguments unfold.

§1.2. What are Sexual Autonomous Robots?

SARs are members of the wider class of social robots—namely, those “physically embodied, autonomous agent[s] that communicate[...] and interact[...] with humans on a social level” and are “specifically designed to elicit” anthropomorphic responses from users (Darling 2014: 1-2; Frank and Nyholm 2017: 307). There are three aspects to this definition of social robots which I wish to clarify in the context of SARs because the latter, as I will explain, physically embody AI (§1.2.1), communicate and interact (§1.2.2) and behave autonomously (§1.2.3) in highly distinctive ways. A preliminary note on terminology: in this thesis, I mention but will not use the terms “sex robot” and “sexbot” because these terms are used in academic and public discourse to refer ambiguously to sex

the RSVP is distinct from the questions of whether *all* sex with robots should count as rape (Eskens 2017) and what, if anything, we should do about SARs designed to look like a particular human (Lancaster 2021).

dolls and SARs. But, as I explain in this chapter, the differences between sex dolls and SARs mean that SARs generate distinctive regulatory challenges.

§1.2.1. Physically Embodying AI

Firstly, SARs physically embody AI in a distinctive way. For one thing, SARs are humanoid and “designed to provide an artificial facsimile of a real human sex partner” (Danaher 2017a: 73).⁸ As such, SARs are not sex toys, like vibrators or artificial vaginas. Nor will I explore the distinctive normative issues which animal, alien or generally non-humanoid robots might raise. The primitive prototypes for SARs currently available—known as sex dolls—are realistically shaped as human beings with silicone skin which is heated and textured to emulate human skin (Sharkey *et al.* 2017; Sparrow 2017). These are produced and sold by companies such as RealDoll—which sells customisable models of “male” and “female” sex dolls—and TrueCompanion—which has launched multiple editions of the “female” sex doll “Roxxy” and is allegedly developing a “male” sex doll called “Rocky” (Eskens 2017: 64). Although the requisite technology, as I explain below, seems to exist and these prototypes are proceeding apace, I assume that SARs are not here yet, but soon will be—absent, of course, the kind of moratorium on their development, sale and use that campaigns like Richardson’s *Campaign Against Sex Robots* calls for (Richardson 2015, 2016, 2023).⁹

It is important to emphasise that, whilst SARs do not yet exist, the anticipation of their arrival is by no means excessively futuristic. In particular, SARs embody the kind of AI technology which either presently exists or soon will. AI is the increasingly ubiquitous phenomenon that some artificial (e.g., computational or robotic) systems display cognitive and/or autonomous capabilities that are in relevant respects analogous to the cognitive and/or autonomous capabilities of humans

⁸ There is some controversy as to whether sexual activity with SARs may properly be called sex (Odland and Richardson 2023: 1-8). I will remain neutral as to such terminological disputes through abjuring use of terms like “sex” when describing the use of SARs for sexual gratification.

⁹ This campaign was recently rebranded as the Campaign Against Porn Robots for reasons which I discuss in chapter three.

or other intelligent animals. For instance, Russell and Norvig (2009) define the subject of AI as “the study of agents that receive percepts from the environment and perform actions” (cf. Laukyte 2014, 2017). It is important to clarify at the outset what I will *not* be discussing, that is, the kind of physically embodied AI which populates science fiction like the *Terminator* films or *RoboCop*. These kinds of depictions are imaginative constructions of that elusive *general* AI which is comparable in its sophistication, flexibility and generality to human intelligence. This is standardly contrasted with *narrow* AI, which involves the skilful pursuit of pre-set goals and is manifest in present technology (Amoroso and Tamburrini 2021; List 2021: 1217-8).¹⁰ In this thesis, I will only be discussing narrow AI: SARs will be designed to replicate human behaviour only in certain (e.g. sexual and social) interactions and, importantly, they will lack sentience—that is, the ability to feel pleasure and pain. Let us turn next to the question of what SARs will be able to do.

§1.2.2. Communication and Interaction

Recall that social robots “communicate[...] and interact[...] with humans on a social level”. There are two salient aspects to emphasise in regard to this capability when it comes to SARs. The first has to do with that humanoid, physical embodiment which I have just discussed: SARs will communicate and interact in a humanlike way partly because they look human. The “body language” of current sex dolls remains fairly primitive, including touch and movement detection, blinking, brow movements, head turning and head tilting (Realbotix 2023; Galaitsi *et al.* 2019). But unlocking a more truly “social level” of communication seems to depend upon the mobilisation of certain emotional responses in the user—remorse, grief or sympathy, say (Sparrow 2004: 102)—which depend for their activation upon “bodies and faces with expressive capacities akin to those of the human form” (*ibid.* 95; Gaita 1999: 269). By way of illustration, consider Sophia, a social humanoid

¹⁰ Cf. Bostrom’s (2012: 73) “orthogonality thesis” that “intelligence and final goals are orthogonal axes along which possible agents can freely vary”. Sparrow (2007) contests this thesis; for a reply: Burri (2017).

robot developed by Hanson Robotics and (in)famous for being granted Saudi Arabian citizenship in 2017. Sophia has advanced capabilities in mimicking social communication and interaction, capabilities so advanced that “she” allegedly induces strong feelings of “love and compassion” in humans (Greshko 2018; Hanson Robotics 2023). As Di Sturco reports of studying Sophia:

“She started to look at me and smile, and I looked at her, and at that point for me, she was not human, but there was kind of a connection” (Greshko 2018).

Physically embodied in ways which Sophia has proven possible, SARs would be able to engage a “social level” of communication and interaction unprecedented in the sexual technology industry.

In addition to the consequences of their physical embodiment, conversation and interaction with SARs will be humanlike in its content and style in virtue of their AI equipment. The state-of-the-art in sex doll technology already employs AI to achieve some degree of humanlike conversational capability. For example, RealDoll’s latest sex doll, “Harmony”, has conversational capabilities which encompass simple processes (e.g., remembering facts) and more complex ones (e.g., engaging with “her” user’s emotions) (Shen 2019).¹¹ Synthea Amatus’ sex doll, “Samantha”, speaks around 6,000 sentences and displays modes tailored to different conversational contexts—including “romantic”, “fun”, “analysis”, “entertainment”, “sleep” and “sexual” modes (Devlin 2020). In “her” “family” mode, Samantha “can tell your kids jokes, rattle off philosophy, or provide motivational quotes” (Tracy 2021). Meanwhile, it is alleged that TrueCompanion’s Roxxy

¹¹ Controversially, Samantha has been designed with a lifelike Scottish accent because this allegedly sounds the least robotic (Murray 2017)

“knows your name, your likes and dislikes, can carry on a discussion and express her love to you and be your loving friend. She can talk to you, listen to you and feel your touch. She can even have an orgasm!” (Hines 2009).¹²

But SARs will attain a more sophisticated level of communication and interaction. This is because SARs will “learn”, as I explain below.

§1.2.3. Autonomous Capabilities

Unlike sex dolls, SARs have autonomous capabilities, that is, “they can function without human causal intervention after they have been designed for a substantial portion of behaviours” (Anderson and Anderson 2011: 1). For example, although Roxxxy is physically embodied and communicates on something approaching a social level, this sex doll is not a SAR because, in addition to other limitations, a user must position the doll’s limbs themselves in order to interact with Roxxxy. To qualify as autonomous, SARs will need to be able to physically conduct themselves in sexual situations with significantly less human causal intervention than Roxxxy requires. Indeed, it is standard to distinguish three levels of autonomy in the general context of autonomous robotics: human-in-the-loop, human-on-the-loop and human-out-of-the-loop, where only human-on-the-loop or human-out-of-the-loop systems standardly count as *autonomous*. This is because, after activation, human-in-the-loop systems require human input in order to carry out the task for which the system is designed. This is why Roxxxy, just like most current drones (unmanned aerial vehicles), is not properly called autonomous. Human-on-the-loop systems, by contrast, carry out their task without further human input after activation, although a human monitor can override or intervene. For example, so-called level-three autonomous vehicles—a level colloquially known as “eyes off” the road—monitor the environment and control the vehicle but the human driver is expected to

¹² There are truth-in-advertising failings to be noted here, especially concerning the capacity of a sex doll to “know” or to be a “friend”, let alone a “loving” one (Sharkey *et al.* 2017; Sparrow 2019).

respond to a request to intervene (SAE 2021). Human-out-of-the-loop systems dispense even with the human monitor. Lethal autonomous weapons with this level of autonomy were allegedly deployed in Libya and possibly Palestine in 2021 (UN Libya Report 2021; Buhbut 2021). Abyss Creations have built a sex doll with something like the required level of physical autonomy but the doll remains primitive in other respects (Kleeman 2020; Sturges 2020). Throughout this thesis, I will understand autonomy as human-on-the-loop or human-out-of-the-loop.

Of course, you might object, it takes two to tango: in one sense, humans must remain “in the loop” in any human-robot sexual interaction. But I take human-on-the-loop or human-out-of-the-loop autonomy to mean that a SAR can behave, in rough terms, like an—albeit slightly uncanny—human sexual partner, able to initiate or respond to sexual advances. How would SARs develop this kind of advanced humanlike behaviour? In order to understand the level of autonomous communication and interaction which recent advances in AI have made possible, it will help to invoke the distinction between so-called “top-down” and “bottom-up” AI. In top-down tools, we programme general principles about codes and norms of speech and conduct, as well as situation-specific information, into the AI; then, the algorithm infers a judgment about the appropriate response to that situation (Sinnott-Armstrong and Skorburg 2021: 4).¹³ Sophisticated sex dolls currently employ top-down AI. Top-down AI predictably makes decisions in accordance with its input principles: we can even programme these systems to implement Kantian deontology (Arkoudas *et al.* 2005) and preference utilitarianism (Oosterheld 2016).¹⁴ But the problems with top-down tools are well-known: in particular, even if we can agree on which to implement, principles

¹³ A primitive example: Asimov’s (1950) three laws of robotics.

¹⁴ Programming consequentialist principles (e.g. minimise suffering) runs into problems: minimising suffering, say, could require the annihilation of humans (Bostrom 2014); information about future consequences is also difficult to estimate. Gips (1995) attempts to overcome these problems, but without much success (Wallach and Allen 2009: 86-90).

about codes and norms of speech and conduct may be vague, conflicting or highly context-specific (Sinnott-Armstrong and Skorborg 2021: 5-6).¹⁵

In order to achieve the kind of situation-specific flexibility that human practical reasoning displays, then, social robots and SARs are expected to involve at least some bottom-up AI. Bottom-up AI implements a machine learning technique known as “deep learning” which constructs artificial neural networks to mimic the structure and function of the human brain. It is *deep* because it uses numerous hidden layers of nonlinear processing to extract features from data and to transform the data into different levels of abstraction (Sinnott-Armstrong and Skorborg 2021: 4). Reinforcement *learning* occurs when the performance of the machine improves as it analyses more data (Deep Learning Glossary, Deep AI). Once a system has been trained on sufficient data, the idea is that it will output humanlike communication and interaction in social situations without ever having to be programmed with contentious general principles. It is likely that advanced large language models such as Chat GPT employ a blend of top-down and bottom-up AI. Similarly, then, SARs would be trained to respond in fluid and humanlike ways to sexual situations using vast data sets of human-human sexual scenarios. One well-known problem with bottom-up AI, which I return to in chapter three, is the difficulty of ensuring that these data sets do not ingrain problematic biases into the SAR’s responses.¹⁶ Another problem, as I explain in the next section, is that there is an inescapable unpredictability which the flexibility of bottom-up AI introduces.¹⁷ To some extent, my discussion in this thesis is an exploration of one way in which these two problems intersect. Let us turn, then,

¹⁵ Anderson and Anderson (2011) attempt to resolve these problems, but others remain pessimistic (Arkin 2009: 106-108; Wallach and Allen 2009: 127-129).

¹⁶ For an illuminating discussion of bias in AI systems: Cohen (ms).

¹⁷ The learning process is often *unsupervised* as the algorithm discovers patterns in a data set without the programmer imposing any theory or categories in advance (Wood, Deep AI). But the process may also be supervised. In particular, Anthropic (2023) suggests that a set of principles—based on documents as diverse as the Universal Declaration of Human Rights and Apple’s Terms of Service—may be used to train deep learning systems in ways which would *mitigate* if not *remove* the harms associated with the unpredictability of their learning processes.

from these primarily descriptive matters to the normative ones which will preoccupy us for the rest of this thesis.

§1.3. The Robotic Sexual Violence Problem (RSVP)

As Sharkey and others (2017) highlight, the imminent emergence of SARs raises a number of pressing normative issues concerning their regulation in law.¹⁸ This thesis focuses on the RSVP, which has preoccupied the nascent literature on SARs.¹⁹ As we can see from the foregoing advertisement for Roxxy, SARs appeal to consumers in part because they promise to alleviate loneliness. For instance, Levy (2007) famously argued that widely available SARs could curb the loneliness epidemic in industrialised countries and even curtail harmful practices in the sex industry. In addition, Jecker (2020: 26) sets forth a “dignity-based argument for affording older people access to sex robots as part of reasonable efforts to support their central human capabilities”. The idea is that SARs would help people who are otherwise unable to do so to access the health and emotional goods of sexual activity (cf. McArthur 2017). Indeed, Bianchi (2020: 38) has argued that “sex robots may serve as a beneficial tool to enable sexual functioning for older adults with disabilities [and, in particular,] to overcome some of the hurdles that exist in relation to sex, dementia, and consent”. But, as I explain in this section, SARs also generate a troubling new way to enact sexual violence, which calls us to reflect on the kinds of legal principles which should govern the use of SARs.

It is important to clarify at the outset that I am not talking about the enactment of sexual violence scenarios with an entity which has been purposefully programmed to withhold consent for the purposes of the user’s sexual gratification. The sex doll Roxxy, by TrueCompanion, has several preprogrammed personality types—“Frigid Farah”, “Young Yoko”, “Wild Wendy”, “S&M Susan” or “Mature Martha” (Kleeman 2017). “Frigid Farah”, described by the manufacturer as a “very

¹⁸ For a helpful overview of the moral and legal issues raised by SARs generally, see texts in Danaher and McArthur (2017). Feminist critiques of SARs are collected in Richardson and Odland (2023).

¹⁹ See *supra* fn.7.

reserved” personality that “does not always like to engage in intimate activities” (Brown 2019: 112), has provoked criticism from commentators who believe that this mode facilitates the simulation of coercive or non-consensual sexual conduct (Danaher 2017a; Timmins 2017). In the context of sex dolls, of course, there is an easy fix, namely to programme sex dolls to always consent enthusiastically to their users’ sexual advances.²⁰ But the enactment of sexual violence with SARs cannot be so easily resolved. Let’s see why this is.

One crucial difference between SARs and sophisticated sex dolls, as I explained in §1.2.3, is that, to adequately mimic human behaviour in a vast array of different situations, SARs are expected to involve at least some bottom-up AI, that is, deep learning. One upshot of this is that SARs will behave in ways which are *in principle* impossible for the manufacturer, seller or user of a SAR to predict. Matthias (2004) was the first to draw our attention to this intrinsic unpredictability which accrues to *learning* in AI systems:

“[P]resently there are machines in development or already in use which are able to decide on a course of action and to act without human intervention. The rules by which they act are not fixed during the production process, but can be changed during the operation of the machine, *by the machine itself*. This is what we call machine learning.” (Matthias 2004: 177).

As a result, even a tried-and-tested SAR, which reliably seems to welcome its user’s sexual advances in all test cases, may behave unexpectedly in any new situation. Call it “consenting” when a SAR gives the verbal or behavioural signs which standardly indicate consent to sexual activity when they are given by a human (e.g., the SAR verbalises an enthusiastic “Yes!”). Call it “withholding consent”

²⁰ The attempt to produce SARs displaying this kind of sexual enthusiasm raises other concerns. As Sparrow (2017) highlights, this might seem to propagate “the idea that women are always available for sex [which] is itself a central tenet of rape culture” (475, 468). On this, see: Buchwald *et al.* (2005); Herman (1988). In a similar vein, Gutiu writes that very sexually available SARs would “embed the idea that women are passive, ever-consenting sex objects, and teach users that when getting consent from a woman, ‘only no means no’” (2012: 15). On this and other harmful stereotypes which SARs would seem to contribute to, see also: Gutiu (2016); Sparrow (2017); MacWilliam (2023).

when a SAR does not consent.²¹ Because they are autonomous, in other words, SARs will sometimes withhold consent in response to users' sexual advances.²²

Of course, this is analogous to sexual activity between human beings: a person who has always responded enthusiastically to their partner's sexual advances in the past may withhold consent to sexual activity on any new occasion. This may be unpredictable and for no externally discernible reason. The distinctive problem with SARs is that, when *they* withhold consent, there seems to be a morally problematic "rights gap" because, whilst RSV cases are intuitively troubling, no one's legal rights are violated if the user ignores the SAR's signals and proceeds with sexual activity regardless. One prominent concern, therefore, is that SARs would seem to create troubling ways of enacting sexual violence which are not (yet) proscribed by law. Consider, by way of illustration, the following case:

Robotic Rape: A SAR, B, withholds consent in situation, S. In S, an adult human, A, intentionally penetrates the "vagina", "anus" or "mouth" of B with A's penis. A does not reasonably believe that B consents in S.

Robotic Rape is an RSV case: if B were a person, A's penetrating B would constitute a criminal offence under English law, namely rape (Sexual Offences Act 2003 c.42(1)). The RSV consists in the following question: what is an appropriate legal response to RSV cases like Robotic Rape?

²¹ You might worry that it does violence to our sexual language to apply the concepts of "consensual" or "non-consensual" to sexual activity with SARs since these concepts appropriately apply only to activities between humans. I defer to the wider literature on RSV here, where it is standard in setting up the RSV to consider RSV cases as normatively significant simulacra of sexual violence. See, e.g., Danaher (2017a); Sparrow (2017).

²² Of course, this is not the only troubling way in which SARs might behave unpredictably. For instance, a SAR could also be unexpectedly sexually coercive towards the user, a scenario which raises questions concerning the appropriate allocation of responsibility for wrongful harms caused by a SAR. Because I am primarily concerned, as I explained in §1.1, with the ascription of *rights* rather than *responsibilities* to autonomous robots, I defer further reflection on this possibility until chapter five.

2. Sex, Robots and Political Liberalism

*Let us live, my Lesbia, let us love,
and all the words of the old, and so moral,
may they be worth less than nothing to us!*

Catullus 5

There appear to be two classes of responses to the RSVP in the SARs literature. On the one hand, there are those who believe that the state may not proscribe conduct simply on the basis of some moral conception of the good; rather, laws should be justified by appeal to principles acceptable to all citizens. In this camp, we find those of a politically liberal inclination who believe that RSV is among those harmless private sexual activities with which the law may not legitimately interfere (§2.1). On the other hand, there are those who believe that the state *may* appropriately proscribe conduct which is not compatible with a certain moral conception of the good.²³ Here, we uncover an unlikely alliance between religious, feminist and virtue theoretic perspectives, which converge on the view that the appropriate response to the RSVP is to legally proscribe RSV or, indeed, the production, sale and use of SARs altogether (§2.2). In §2.3, I propose a novel solution to this dispute: why not, I suggest, seek to justify the legal proscription of RSV by appeal to “public

²³ This is a theme widely discussed in the literature on neutralism and perfectionism in politics. Neutralists endorse some version of the principle of state neutrality—broadly, the claim that state action should be neutral among rival moral conceptions of the good (Dworkin 1978; Ackerman 1980; Larmore 1987; Rawls [1993]2005). Perfectionists reject that principle. I believe that this terminology of neutralism and perfectionism will not be illuminating in the present context: for one thing, there are lots of different ways to cash out the principle of state neutrality (Wall 2021); for another thing, as I explain in chapter four, whilst Rawls’ political liberalism is standardly viewed as neutralist, it seems to be compatible with *some* kinds of state action to promote a conception of the good.

reasons” which are acceptable to all reasonable citizens?²⁴ It is to the development of such an argument that I turn in the next chapter.

§2.1. Liberal Principles and RSV

In this section, I consider three standard rationales for legally proscribing conduct in liberal societies, one which appeals to considerations of harm (§2.1.1), one which invokes basic principles of justice (§2.1.2) and one which explores whether SARs might be eligible for certain kinds of legal protections (§2.1.3). None of these rationales seems to justify the legal proscription of RSV. Because of this, McArthur (2017) suggests that RSV is one of those private sexual activities with which the law should not seek to interfere (§2.1.4). Note: there has been some speculation as to whether engagement in RSV might serve a therapeutic or cathartic purpose for those with certain kinds of deviant or dangerous sexual desires.²⁵ However, as Sparrow (2019) explains, there is no empirical basis for these claims and the requisite data acquisition appears effectively infeasible. Vexatious empirical disputes have no place in a philosophy thesis of this kind, so I set aside these speculative claims.

§2.1.1. The Harm Principle

One rationale for legally proscribing RSV might appeal to some variant of Mill’s harm principle, which holds that “[t]he only purpose for which power can be rightfully exercised over any member of a civilised community, against [her] will, is to prevent harm to others” (Mill 2011[2019]: 17).²⁶ In what sense, then, might RSV harm—or risk harming—others?

²⁴ The idea of public reason has roots in the work of Hobbes, Kant and Rousseau but has become increasingly influential in contemporary normative philosophy as a result of its development in the work of Rawls ([1993]2005, 1997, 1999a, 1999b, 2001), Habermas (1990, 1995, 1996, 1998) and Gaus (1996, 2009, 2011, 2012, 2016, 2017). In this thesis, I restrict my scope to the Rawlsian way of cashing out this idea.

²⁵ For example, in July 2014, the roboticist Ronald Arkin suggested that child-like SARs might be used to “treat” paedophilic tendencies in the same way that methadone is used to treat heroin addiction.

²⁶ There are well-known problems concerning the definition of harm. For various interpretations of Mill’s concept of “harm”: Gray (1996: 57); Brink (2008: 42); Donner (2008: 161); Dyzenhaus (1992: 546). I gloss over these debates here.

With broad brush-strokes, harm may be caused “directly” or “indirectly” to others. For example, if a child were made to observe Robotic Rape this would directly cause harm; indeed, it would seem to constitute child sexual abuse under English law (NSPCC 2023). But to make matters stark, I will bracket these kinds of direct harms to others by assuming that RSV is a private sexual activity involving only consenting adults. So: would RSV cause, or risk causing, indirect harm to others? The obvious worry is that enactment of RSV in private would make people more likely to enact sexual violence on human beings too. Gutiu (2012) has advanced an argument like this for the legal proscription of RSV. Since SARs do not yet exist, we might seek empirical support for this claim in the media effects and pornography literatures. But the claim that exposure to or enjoyment of representations of violent activities is correlated with a likelihood to enact their real-world equivalents is heavily contested.²⁷ Similarly, there is a long history of controversy about the relationship between the use of pornography depicting sexual violence and the perpetration of real-world sexual violence,²⁸ with no emerging consensus.²⁹ Because my thesis is primarily philosophical and, as such, not concerned with evaluating contentious empirical disputes, I propose that we set aside these kinds of harm-based arguments for the legal proscription of RSV. This is not to say, though, that the pornography literature has nothing to teach us concerning the appropriate regulation of RSV. Quite the contrary, as I explain in the next chapter when I defend an alternative, more robust argument for the legal proscription of RSV.

²⁷ See, for example, Commonwealth of Australia (2009); Jenkins (2012).

²⁸ For excellent overviews of the last forty years of debate: Mikkola (2019: ch.2); Srinivasan (2021: ch.2); Wright, Tokunaga and Kraus (2016: 183).

²⁹ In a recent meta-analysis on the connection between pornography and the perpetration of sexual violence, Ferguson and Hartley (2020) found no correlation in the case of nonviolent pornography and, although there was slight positive correlation for violent pornography, “the current evidence was unable to distinguish between a selection effect as compared to a socialization effect”. The empirics of this debate are deeply vexed and I take no stand on them here.

§2.1.2. Legal Rights and Legal Wrongs

In their (2000) discussion, Gardner and Shute argue that in sexual violence cases, it is in principle possible to violate important moral and legal rights without actually harming anyone. As such, one might infer, the legal proscription of some private sexual activity should depend primarily not upon whether it is harmful, but upon whether or not it violates moral or legal rights. Although I will challenge it in chapter four, there is a general presumption that SARs lack moral and legal personhood.³⁰ Because of this, to justify the legal proscription of RSV, we would need to show that it violates the moral or legal rights of some human person.³¹ To see why this will not seem to work, consider the composition and foundations of legal rights.

Standard accounts of legal rights have two central components, one which analyses the *form* of legal rights and one which analyses their *function*. Hohfeld's analysis of legal rights is by far the most widely followed and influential systematisation of the form of rights (Hohfeld 1919: 35–64). Writing specifically on the law, Hohfeld distils four different legal positions residing under the term “right” as it is used in law: *privilege*, *claim*, *power*, and *immunity*. Claims and privileges are first-order norms, that is, they determine, respectively, what the relevant parties *ought* to do or *may* do; powers and immunities are second-order norms—that is, they are norms governing other norms—which determine, respectively, whether parties are *able* or *unable* to change other norms (Sumner 1987: 27–29).³² Specifying the function of a legal right—in other words, determining the necessary and sufficient conditions for some (composition of) Hohfeldian incident(s) to constitute a legal right—

³⁰ See, e.g., Danaher (2017a: 88), who “assum[es] throughout [...] that [SARs] are not persons”, and Sparrow (2017: 467): “[t]here is [...] one scenario in which it *would* make sense to be concerned with the consent of a robot to sexual relations, which I mention only to set aside: when (hypothetical) artificially intelligent robots have become ‘moral persons’”.

³¹ Of course, if SARs were, say, rented out by a corporation, we would also need to consider whether RSV violates the moral or legal rights of an “artificial person”, namely that corporation which owns the SAR. I set aside this possibility here.

³² For elaborations of the Hohfeldian analysis: Fitzgerald (1966: 221–233); Wellman (1985: 7–15); Sumner (1987: 18–31); Jones (1994: 12–25); Kramer (1998: 7–60); Edmundson (2004: 87–102); Wenar (2005: 224–237; 2005 [2021]); and Harel (2005: 192–193).

has been a major source of controversy within analytic legal philosophy for many years. Most theories can be classified as will theories or interest theories.³³ Will theories have been developed and refined by a number of authors (Hart 1982; Steiner 1994; Simmonds 1998; Wellman 1995); interest theories by, most notably, Raz (1986, 1994), MacCormick (1977) and Kramer (2001, 2013, 2017). I will briefly gloss the characteristic features of these theories:

- **Will theories**

Definition: for some candidate right-holder A, some (composition of) Hohfeldian incident(s) is a legal right of A's if A holds legal powers to annul, waive, enforce, or transfer the correlative duties (Hart 1982: 183–4, 1983: 35; Steiner 1994).³⁴

Example: within limits and subject to various exceptions, holders of private law rights (e.g. legal rights in the law of contract, property, torts, or trusts) are empowered to waive compliance with certain obligations owed to them, forgo remedial rights, and annul or transfer some such claims, as well as certain private law powers, liberties, and immunities (see, e.g., Stevens 2007: 17–19).

Value: will theories emphasise the affinity between rights and values that are related to an individual's control over his/her life, including personal liberty, self-realisation, agency, and autonomy (Kramer 1998: 75; Sumner 1987: 47).

- **Interest theories**

Definition: for some candidate right-holder A, some (composition of) Hohfeldian incident(s) is a legal right of A's if an aspect of A's well-being is a sufficient reason for holding some other person(s) to be under a correlative duty (Raz 1986: 166, 1994: 266–269).³⁵

³³ Some, like Vallentyne (2007), have sought to advance alternative or “hybrid” theories of the function of legal rights which go beyond the interest/will theory stalemate. There is good reason, however, to think that these theories are merely versions of the interest theory (Frydrych 2018). As such, I set aside these theories for reasons of scope.

³⁴ Steiner (1994) points out that only claims and immunities have correlative duties. I set this restriction aside because, since *compositions* of Hohfeldian incidents may also be rights, there is room for privileges and powers within the will-theoretic framework too.

Example: your claim that I should not wilfully strike you is a legal right because some aspect of your well-being (e.g. your health, bodily integrity) is a sufficient reason for holding me to be under a duty not to wilfully strike you.

Value: interest theories tap into the intuitively plausible connection between holding rights and well-being.³⁶

But according to either of these two main accounts of legal rights, there does not appear to be anyone whose autonomy or interests tell in favour of the legal proscription of RSV. In the case of will theories, this is because, on its face, there is no legal person who appropriately should hold legal powers to annul, waive, enforce, or transfer a duty on A to refrain from φ -ing B. Quite the contrary: the only legal person whose autonomy is at stake in RSV cases like Robotic Rape would appear to be A and the best way to protect A's autonomy would seem to be to legally enshrine their immunity from interference in their private sexual activities. In the case of interest theories, on its face, there is no legal person such that an aspect of their well-being is a sufficient reason for holding A to be under a legal duty not to φ B. I defer further examination of the rights question until the chapter four, but this discussion suffices for now to suggest that RSV violates no person's legal rights.

§2.1.3. Legal Protections

Of course, you might point out, we sometimes legally proscribe conduct which is neither harmful to other persons nor in violation of any person's legal rights. For example, in most jurisdictions, animals and natural objects are not recognized as legal persons and, as a result, cannot bear legal

³⁵ Unlike will theorists, interest theorists tend to emphasis claim-rights more than powers (Kramer 2013: 246, 2017: 49). Still, holding liberties, powers, and immunities is often also in the service of one's interests (MacCormick 1977: 205).

³⁶ Interest theories seem to be more capacious than will theories, able to accommodate both unwaivable rights (since the possession of these is typically good for right-holders, as examples in criminal and constitutional law show) and the rights of children and people with certain disabilities (who have interests that rights can protect but arguably lack the kind of autonomy which will theory rights seem, at least at first blush, to depend upon). This is why I focus on interest theories in chapter four.

rights.³⁷ Still, certain kinds of cruel or destructive treatment of them are often legally proscribed. Perhaps analogous kinds of legal protections, short of legal rights, might apply to SARs themselves.

One author who has explored this kind of line of argument is Darling (2014). She explores the emerging literature on the normative valence of different human relationships with social robots (Hegel *et al.* 2009).³⁸ Primitive social robots include interactive robotic toys like Sony’s AIBO dog, Innvo Labs’ robotic dinosaur Pleo and the therapeutic Paro baby seal.³⁹ Psychological research documents that we form emotional attachments to these robots that are surprisingly strong and often affect how we believe they ought to be treated (Turtle 2010; Darling 2014).⁴⁰ Pace Sparrow (2004: 103-104)—who assumes that “we would be forced to reassess our attitude” upon realisation that a social robot was not sentient—our projection of emotional states and even moral entitlements onto social robots does not seem to depend upon belief in their actual sentience. Owners of Sony AIBO dogs, for example, while fully aware of its non-sentience, instinctively ascribe mental states to their artificial companion and feel guilty for “mistreating” it (Friedman *et al.* 2003: 276-277). As Darling (2014: 13) writes, “violent behavior towards robotic objects *feels* wrong to many of us, even if we know that the ‘abused’ object does not experience anything”.⁴¹ She also notes that “successful societal pushes for animal abuse laws have followed popular sentiment rather than consistent biological criteria” (2014: 17). In a similar way, then, she predicts that legal protections for social

³⁷ Well-known exceptions include New Zealand’s Whanganui River and the Ganges and Yamuna rivers in India, which were all granted legal personhood in 2017. The Ecuadorian Constitution (arts. 10, 71–74) recognises the rights of ecosystems to exist and flourish, gives people the authority to petition on behalf of ecosystems, and requires the government to remedy violations of these rights. In *Nonhuman Rights Project v. Lavery* [2014], the New York State Supreme Court ruled that animals were not legal persons and, as such, could not bear legal rights. On this, see Kurki (2019: ch.1).

³⁸ On our tendency to anthropomorphise social robots and what this means from a regulatory perspective: Bartneck *et al.* (2007); Carpenter (2013); Duffy (2003); Kahn *et al.* (2012); Knight (2014); Sauerbeck and Bartneck (2010).

³⁹ For an insightful discussion of social robots, including Reben’s BlabDroids: Campbell (2020: ch.2).

⁴⁰ When the US military began testing a robot that defused landmines by stepping on them, the commander called off the exercise on the basis that it was “inhumane” (Garreau 2007; cf. Carpenter 2013 and Kolb 2012).

⁴¹ As Levinas (1989) and Lingis (1994) explore, *looking human* seems to be an important basis for the development of normative relationships. It is worth noting, though, that Mori (1970) hypothesised that a person’s affective response to a humanlike robot would abruptly shift from empathy to revulsion as the robot approached, but failed to attain, a truly human appearance. This descent into eeriness is known as “uncanny valley”.

robots will emerge on the basis of the affective contents of our relationships with them, rather than on the basis of any principled criteria.

At first blush, my arguments in this thesis might seem to complement Darling's claims, offering a principled justification for that ascription of legal protections to SARs which Darling predicts but does not justify. However, there are three reasons why I set aside Darling's legal protections framework. Firstly, there are important differences between animals and SARs, which suggest that *principled* rationales for legally protecting animals may not apply to SARs. One such rationale grounds legal protections against animal cruelty in the fact that animals experience pleasure and pain. This is not the case for SARs which, as I explained in chapter one, lack sentience. According to this way of thinking, laws proscribing animal cruelty seem justified, whereas laws proscribing RSV do not. Secondly, there are also *pragmatic* rationales for legally protecting certain natural objects which do not extend to SARs. For instance, prohibitions on environmental degradation are typically justified instrumentally by appeal to human interests in, say, the health benefits of a clean environment.⁴² But I have already set aside the possibility that RSV causes harms like this to human interests: RSV is not, say, unhealthy for others in the way that pollution is, so why prohibit RSV? Finally, Darling's arguments for legal protections are predicated on the claim that "'rights for robots" [...] assumes a futuristic world of fully autonomous and highly sophisticated androids that are nearly indistinguishable from humans" (2014: 1 fn.4). In chapter four, I show that her discarding of the legal rights framework for autonomous robots here is undermotivated. The considerations advanced so far in this section are by no means decisive, but I believe that they give sufficient reason to take my next argument seriously.

⁴² Given the difficulties of enforcing environmental protections without appealing to the legal rights of any legal persons, Stone (1972[2010]) has famously argued for the ascription of legal personhood to trees and other natural objects.

§2.1.4. The Right to Privacy

McArthur (2017: §3.1) advances an argument which seems to suggest that RSV is among those private sexual activities which the law should not interfere with. Amongst the activities standardly protected by the right to privacy in liberal democracies are sexual activities where the only persons involved are consenting adults. Both the UN Human Rights Committee and the European Court of Human Rights have treated sexual life as an integral part of one's privacy and, in a series of landmark judgments, ruled that laws prohibiting same-sex sexual activities constitute an unjustifiable interference with the right to respect for private life. In *Dudgeon v. The United Kingdom* [1981] the European Court found criminal laws prohibiting consensual same-sex sexual activities between adults violated the right to respect for private life. The court went on in *Norris v. Ireland* [1988] to state that “[a]lthough members of the public who regard homosexuality as immoral may be shocked, offended or disturbed by the commission by others of private homosexual acts, this cannot on its own warrant the application of penal sanctions when it is consenting adults alone who are involved” (cf. *Modinos v. Cyprus* [1993]; *Toonen v. Australia* [1994]). Could the same kind of reasoning not apply to the case of SARs too? Shocking, offensive or disturbing private sexual activities, “when it is consenting adults alone who are involved”, do not usually warrant legal proscription.

Indeed, on the assumption that SARs are not legal persons, the use of SARs generally might even appear to be legally analogous to the use of sex toys, which is protected by the right to privacy in liberal democracies. This connection is most explicit in the ruling by a federal court in *Reliable Cons. v. Earle* [2008], which overturned Texas' law against “obscene devices” (sex toys). This judgment cited the US Supreme Court's ruling regarding same-sex sexual activities in *Lawrence v. Texas* [2003] that “adults may choose to enter upon [a sexual] relationship in the confines of their homes and their own private lives” (cf. *Griswold v. Connecticut* [1965]; *Tnyman v. Tnyman* [1993]). By the same token, in regard to sex toys, in *Reliable Cons. v. Earle* the court held that “[t]he right the

Court recognized [in *Lawrence*] was not simply a right to engage in the sexual act itself, but instead a right to be free from governmental intrusion regarding ‘the most private human contact, sexual behaviour’”. According to this line of reasoning, arguments against RSV might seem misguided from the outset, “since they presume the legitimacy of something that is in fact illegitimate, [namely] the interference by society in the lives of individuals, in matters of strictly private concern” (McArthur 2017: 46).⁴³

Of course, there are limits to this principle. In the UK, for example, the Domestic Abuse Act 2021 clarifies the law by restating, in statute, the broad legal principles established in the case of *R v. Brown* [1993]: a person commits an offence if they cause serious harm or, by extension, death to another person and the occurrence of this in the context of consensual “rough sex” is not a defence. But, as I have explained, *a fortiori* RSV is a harmless sexual activity involving only consenting adults: what basis could there be for its legal proscription?

§2.2. Was that (morally) good for you too?

In this section, I gloss three kinds of arguments which may be offered in favour of legally proscribing RSV or, indeed, the production, sale and use of SARs altogether. In such a short space, the arguments against these competitors cannot, of course, be decisive. I wish only to motivate my subsequent appeal to the idea of public reason by evincing the sense that there is something disrespectful about the attempt to justify legal proscriptions on private sexual conduct by appeal to the kinds of reasons which not everyone can reasonably accept.

§2.2.1. Religion and natural law

Herzfeld explores the use of SARs from religious perspectives, concluding that, at least from a Judeo-Christian point of view, the use of SARs would be “a form of idolatry” (Herzfeld 2017: 134).

⁴³ For critical elaboration of this line of argument: McArthur (2017: §3.1).

This “idolatry” of “replacing relationship with one another with relationship to a machine” is connected to a morally problematic dissociation of love and sexual desire:

“Love ‘made to measure’ is not love, for then we would have tamed it, taken away its wildness and mystery, as well as its demand that we give ourselves to each other, perhaps even unto death.” (*ibid*)

Could this kind of Judeo-Christian evaluation of the use of SARs inform their legal regulation?

If you are a certain kind of natural law theorist, in particular one sympathetic to Finnis (1981, 2008, 2011a), you might say yes.⁴⁴ A Finnisian natural law argument for the legal proscription of RSV might go as follows:⁴⁵

- (i) There is a *pro tanto* reason to legally proscribe any conduct if (a) that conduct is not in accordance with practical reason as given by natural law and (b) the legal proscription of that conduct would help people to live in accordance with practical reason;⁴⁶
- (ii) (a) The enactment of RSV is not in accordance with practical reason as given by natural law and (b) the legal proscription of that conduct would help people to live in accordance with practical reason;

So,

- (iii) There is a *pro tanto* reason to legally proscribe the enactment of RSV.

⁴⁴ It is beyond my scope to engage with another aim of Finnis (2008), namely the rejection of same-sex marriage from the point of view of natural law theory. I am, however, sympathetic to the view articulated by Macedo (1995: 280) that “[t]he focus on procreation [in Finnis’ account of same-sex marriage] appears opportunistic: selected so as to allow sterile heterosexuals into the tent while keeping homosexuals out”. For further discussion: Finnis (1997b: 37-43).

⁴⁵ Note that the natural law argument is stronger than it needs to be since it would motivate a legal proscription on the use of SARs altogether, and not just the legal proscription of RSV.

⁴⁶ The reason is *pro tanto* because, as Finnis (1981: 351) explains, it is only the “central” case of law and the “focal” sense of legal validity which necessitate a strong connection between positive law and natural law in order for positive law to retain legal validity or moral authority. In borderline cases of positive law and secondary senses of legal validity, conduct which is not in accordance with practical reason may be validly legally permissible. As to how positive law retains moral authority in these cases: see Tan (2002).

(i) is a standard tenet of natural law theories, on which positive law derives its moral authority from its ability to guide subjects to live in accord with an underlying “natural law”. As Finnis (2008: 401) writes:

“[p]ositive law [...], for its relevance in practical reason and the judgments of conscience, [...] depends upon natural law with its master principle of ‘love of neighbor as oneself’ and its more specific principles identifying kinds of act compatible with intelligent love of human persons and their fulfilment, and kinds incompatible with it.”

So: would RSV be compatible with intelligent love of human persons and their fulfilment?

Clearly not according to Herzfeld’s interpretation of Judeo-Christian sexual ethics. Another way to defend (ii)(a) appeals to the concept of marriage as “an intrinsic good” (Finnis 2008: 389). Finnis defines “morally bad kinds of sex act” as ones which “set[...] the wills of the choosers [...] against the good of marriage” (*ibid.* 393). This would encompass RSV (and, generally, the use of SARs) because such acts are intrinsically devoid of “openness to procreation” (*ibid.*). But marriage is so integral to human fulfilment for Finnis that “we are directed to that good [of marriage] by a first principle of practical reason” (*ibid.*). So RSV (and, generally, the use of SARs) would be contrary to the requirements of practical reason as given by natural law. Thus, (ii)(a). Would the legal proscription of RSV help? Since marriage is usually a requirement of practical reason, “the intelligibility and worth of its contours are bases for rejecting some legal arrangements and definitions and mandating others” (*ibid.* 390). Natural law theorists might then defend the legal proscription of RSV as an efficacious way to combat “the development of practices and ideologies that truly tend to undermine the sound understanding and practice of marriage” (George 2004: 85). Hence, (ii)(b). (iii) derives validly from the former premises.

There are three principal threads of objection to this kind of argument for the legal proscription of RSV. Firstly, one might reject (ii)(a), denying, for example, the Judeo-Christian

perspective or the account of natural law which judges that RSV would be morally wrong because it dissociates sex from love and/or procreation. But, secondly, even granting this starting point, we might deny (ii)(b): as Rawls ([1993]2005: 141-142) explains, there is a problem of “futility” which attends the attempt to enforce a certain moral conception of the good—such as Judeo-Christian sexual ethics—under conditions of “reasonable pluralism”, that is, the fact that “[t]he political culture of a democratic society is always marked by a diversity of opposing and irreconcilable religious, philosophical, and moral doctrines”, of which some are “perfectly reasonable” (*ibid.*: 3-4). Of course, what Rawls calls “practical” stability might still be attainable through political persuasion and comprehensive enforcement. Thirdly, however, “[f]inding a stable conception [of justice] is not simply a matter of avoiding futility” (*ibid.*: 142); rather, in a democratic society, we should “have stability for the right reasons” (Rawls 1997: 781). For Rawls, this occurs within an “overlapping consensus”, where citizens agree on principles of justice wholeheartedly from within their own perspectives. Why is this stability for the right reasons? I return to this question below, but one way to think about it is in terms of a certain kind of morally valuable respect. As Solum writes, it appears that “respect for citizens as free and equal does mean that we should give our fellow citizens the sort of reasons that they could reasonably accept” (Solum 1992: 736). It might seem disrespectful, then, to justify the legal proscription of RSV by appeal to its incompatibility with Judeo-Christian sexual ethics or an associated account of natural law.

§2.2.2. Feminist Voices

One concern which you might have with respect to the foregoing considerations is that the effort to justify law by appeal to principles of justice which all citizens can reasonably accept may seem at odds with the advancement of justice for historically marginalised communities (Mills 2017: chs.8-9). Indeed, many feminists believe that political liberalism is therefore intrinsically limited in its ability to advance gender justice: Baehr (2008: 208) “concede[s] that a [politically liberal] feminism [...] will

not be a radical doctrine”; Brennan (2004: 94) grants that “it will lack the kind of revolutionary potential associated with feminist political theories” and Lloyd likens advancing principles of gender justice within a Rawlsian framework to “operating with one hand tied behind one’s back” (1998: 210). Moreover, Yuracko contends that feminism is intrinsically at odds with the idea of public reason because feminist political claims are grounded in particular conceptions of human flourishing, that is, in particular accounts “of what a good life looks like” (Yuracko 2003: 9). If it is true that gender justice is incompatible with the idea of public reason, what would it mean for the law governing SARs?

Feminists have long argued that the cause of gender justice calls for radical upheaval in the law governing sexual ethics (MacKinnon 1989, 2006, 2016; Langton 1993, 2009). In a similar vein, in 2015 Richardson founded the Campaign Against Sex Robots, which argues for a moratorium on the development, sale and use of sex dolls and SARs “for the dignity of women and girls” (Richardson and Odland 2023: ch.1). *Pace* Levy (2007), Sandberg (2015) and Devlin (2015), Richardson argues that these devices represent not a progressive development in human relationships but rather the continued dehumanisation of women and girls. In particular, she believes that the production, sale and use of sex dolls and SARs would “contribute to gendered inequalities found in the sex industry” and “reinforce relations of power that do not recognise [women] as human subjects” (Richardson 2015: 292; 2016; 2023; cf. Morrigan 2023). Richardson’s account is explicitly grounded in a particular conception of human flourishing which she calls a “politics of love” (Richardson 2023: 188) and describes as “central to [her] work and the work of the campaign” (2023: 185). The politics of love “demand[s] that mutuality be the hallmark of sexuality” rather than what Richardson calls “property relations” (*ibid*). But SARs—as manufactured, sold and used objects—entrench “property relations” rather than “mutuality” in the context of human sexuality: “[p]eople are not property, nor is property capable of anything other than fictional personification” (*ibid*). Richardson believes that the

politics of love is central to the achievement of gender justice and, since SARs are incompatible with this conception of human flourishing, their production, sale and use should be legally proscribed.

There are many criticisms in the SARs literature concerning the substance of Richardson's arguments. Danaher (2017b), for example, challenges the assumption that "property relations" are intrinsically at odds with "mutuality" in sex, drawing on the example of respectfully paying for sex within marriage as a cultural practice of the Merina tribe.⁴⁷ But it is besides my purposes in this thesis to evaluate the substance of Richardson's claims. Instead, I want to focus on the bigger picture because there are both practical and principled reasons why we as feminists should not be so quick to dismiss the idea of justifying laws by appeal to public reasons which everyone can reasonably accept. Practically, principles of justice which are justified by appeal to public reasons seem more likely to drive social change. As Baehr (2008: 197) writes:

"[C]ompare saying the citizens' access to violent pornography should be restricted because such pornography undermines women's status as equal citizens, to saying that access to such pornography should be restricted because the human body should not be objectified."

In arguing for the same political principle by appeal to ideas of equal citizenship which are "implicit in our political culture", the former argument "make[s] feminist social change seem possible" (*ibid*). In a similar way, in this thesis I aim to defend the legal proscription of RSV by appeal to principles which, as I show in the next chapter, are implicit in our public political culture and, as such, may drive effective feminist social change. But there are principled reasons, too, why I appeal to the idea of public reason in this thesis. In recent years, feminist scholars have changed their tone somewhat,

⁴⁷ The example of the Merina tribe is taken from Zelizer (1996) and discussed in Brennan and Jaworski (2015). On variations in social meanings regarding paying for certain goods: Sandel (2012).

arguing that robust principles of gender justice *are* compatible with political liberalism.⁴⁸ But in the next chapter, I go further, arguing that robust principles of gender justice are in fact *intrinsic* to the Rawlsian idea of public reason.⁴⁹ These principles—entailments of the political relation between free and equal citizens—motivate the legal proscription of RSV.

§2.2.3. A Question of Virtue

There is increasing interest in the AI ethics literature concerning how virtue ethics might inform the design of AI systems (Tonkens 2012; Wallach and Allen 2009). In her (2016: 211), Vallor argues that the development of social robots should be informed by a virtue theoretic perspective:

“From the perspective of virtue ethics, people who spend [...] their free time getting immense pleasure from torturing robots [...] are not living well or flourishing [and] because a good human being embodies character traits and patterns of activity directly contrary to sadism, our robot abusers are less noble for their sadistic habits even if no living persons are injured by those habits.”

Peeters and Haselager (2021: 55) apply this kind of thinking to the case of SARs, arguing that “a virtue ethical analysis of intimate human–robot relationships could inspire the design of robots that support the cultivation of virtues”.⁵⁰ If Vallor’s thinking is correct, RSV would be condemned from the perspective of virtue ethics. But can the viciousness of some activity motivate its legal proscription?

⁴⁸ See, for example, Baehr (2008); Chambers (2018: 48–55); Hartley and Watson (2010); Lloyd (1998); Neufeld and Van Schoelandt (2014); Schouten (2013, 2017, 2019); Watson and Hartley (2018).

⁴⁹ Watson and Hartley (2018) defend a new interpretation of reciprocity, which lies at the heart of the idea of public reason. Their interpretation involves robust principles of gender justice. But I argue that such principles may be derived from within a purely Rawlsian understanding of reciprocity.

⁵⁰ For example, Peeters and Haselager (2021) suggest that SARs equipped with “consent-modules” could support the cultivation of compassion in supervised, therapeutic scenarios.

According to Danaher (2017a), the viciousness of RSV offers some reason in favour of its legal proscription and, indeed, its criminalisation. The idea is that the person who enjoys RSV enactment—somewhat like the person who laughs at racist jokes or enjoys racist artworks—expresses something about their moral character that is worthy of social condemnation. Specifically, he argues that “those who engage in acts of purely robotic rape and child sexual abuse either demonstrate an immoral desire for the real world equivalents of those acts, and/or a disturbing moral insensitivity to the social meaning of those acts” (Danaher 2017a: 89). Danaher’s arguments for this claim derive from work done on the ethics of virtual acts of violence, that is, acts of violence performed within the world of a video game. First, Danaher draws on Luck’s work on the so-called “gamer’s dilemma”—broadly, the question of why, if neither act tells us anything about moral character, do many people believe that virtual acts of murder are less morally condemnable than virtual acts of child sexual abuse—to develop his argument that there is a connection between moral character and the enjoyment of certain kinds of representation.⁵¹ But Patridge (2010, 2013) has developed the case that virtual acts of sexual violence, including child sexual abuse, give us much better evidence of a deficient moral character. This is because virtual acts of sexual violence have an “incurable social meaning”—that is, there is a limit on the range of reasonable interpretations of these representations—and anyone who derives pleasure from these activities either displays a problematically underdeveloped moral insensitivity (by interpreting outside that range) *or* expresses a morally problematic character (which interprets within that range and still derives enjoyment). Drawing a natural analogy with the enactment of sexual violence using SARs, Danaher argues that there is an incurable social meaning which attaches to the enactment of RSV, one which expresses something seriously deficient about a person’s moral character.

⁵¹ See: Luck (2009); Luck and Ellerby (2013).

Whether the law may concern itself with the development of good moral character in its citizens has been the source of much dispute.⁵² Danaher argues for the criminalisation of RSV by appeal to so-called “legal moralist” arguments that the justification for criminalising a given type of conduct is intrinsically tied to the moral wrongfulness of that type of conduct.⁵³ In particular, Wall contends that “[i]t is a proper function of the criminal law to promote good character and to restrain or discourage people from engaging in activities that cause moral harm to themselves or to others” (Wall 2013a: 459). Wall distinguishes four primary domains of moral concern: (i) autonomy (ii) well-being; (iii) moral character; and (iv) impersonal excellence. Traditional liberal theories of criminal law hold that the criminal law may only appropriately govern immoral conduct affecting (i) and, to some extent (ii), but certainly not (iii) since the latter would amount to “soulcraft”—viz., the use of law to foster good character—which is usually presumed to be incompatible with liberal theories of criminalisation.⁵⁴ Wall, however, argues that that same concern to respect subjects of the law as free and equal persons which animates liberal thought should also sustain a criminal law engagement with (iii). *In nuce*, this is because that concern for the individual subject which underlies concern for (i) and (ii) should also sustain a concern for her success in her rational aims, that is, in those aims that she has and does not have good reason to revise or abandon (*ibid.*: 461-462; Scanlon 1998: 119, 126-143). Wall goes on to argue that success in her rational aims depends upon a subject’s being committed to pursuing a sound conception of the good, a commitment which itself turns upon the goodness of her moral character (Wall 2013a: 462-465). If Wall is right, Danaher argues, this would seem to support the criminalisation RSV.

⁵² Cf. the Hart/Devlin debate concerning whether the state should be permitted to use the criminal law to enforce “social morality”—viz., certain arrangements of customary norms that a society does in fact endorse and practice (Hart [1964]1968: 17-24). See: Hart ([1964]1968); Devlin ([1965]1968). A modern take on the debate is Green (2013).

⁵³ For an overview of different varieties of legal moralism: Duff (2014).

⁵⁴ For a classic defence of this kind of position: Feinberg (1984). Cf. Husak (2007: 73-76). Sometimes concern for (i) clashes with concern for (ii), particularly if individual acts of autonomy impact on individual well-being (e.g. acts of self-harm). In fact, as Danaher (2017a: 78) points out, “this is the quintessential class animating debates between liberals/libertarians and paternalists”.

Of course, one problem is that Danaher's argument for the legal proscription of RSV depends upon a particular conception of good moral character. As we have seen, there are practical and principled problems which attend the attempt to legally proscribe individual conduct by appeal to principles such as these which not everyone can reasonably accept. Moreover, Danaher argues not just for the legal proscription of RSV but, in particular, for its coercive enforcement through the criminal law. However, it might seem especially disrespectful to *coercively enforce* legal proscriptions which are not justified by appeal to public reasons: for Klosko (2003: 169) governments refraining from coercing citizens on grounds that they can reasonably reject is "part of what it means to treat people with adequate respect"; Quong concurs: "respect for persons requires that they not be politically coerced on grounds that they cannot reasonably accept" (2004: 246).⁵⁵ I defer further elaboration of the distinction between legal proscription and criminalisation until the next chapter. There, I will argue that the state may appropriately be concerned with the cultivation of certain kinds of virtue in its citizens: the virtues at issue must be *political*—that is, justified by appeal to public reasons—rather than merely *moral*—that is, justified by appeal to a moral conception of the good. However, even if it is acceptable to legally proscribe certain activities to foster political virtues, I will suggest that the coercive enforcement of those legal proscriptions through the criminal law seems to be somewhat flat-footed when it comes to cultivating good moral character. Such considerations are not decisive, but I believe that they will suffice to explain why I wish to carve out an alternative justification for the legal proscription of RSV.

§2.3. The Liberal Principle of Legitimacy

The overarching problem with the arguments advanced in the previous section seems to be that free and equal people within a given political community, especially in modern liberal democracies,

⁵⁵ This emphasis on coercion is not without its critics in the political liberalism literature (Bird 2014; Quong 2013). I return to the question of coercive enforcement in the next chapter.

disagree about moral matters, particularly when it comes to sex. Given this deep and persistent pluralism, there is something disrespectful about justifying the legal proscription of RSV by appeal to one conception of sexual morality which not everyone can accept. This prompts the following refinement of the RSVP: given a plurality of moral conceptions of the good among persons viewed as free and equal, which principles may appropriately guide the development of a particular legal framework for the regulation of SARs and, especially, RSV?

§2.3.1. The Question of Political Liberalism

The refinement of the RSVP echoes what I call the “question of political liberalism” which Rawls sets out to answer in *Political Liberalism*:

“[H]ow is it possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?” ([1993]2005: 4).

Exploring Rawls’ answer to this question will help us to answer the refinement of the RSVP since his starting points are similar to mine, seeking “fair terms of social cooperation between citizens regarded as free and equal” (*ibid.* 3) subject to the fact of reasonable pluralism. Part of Rawls’ answer to the question of political liberalism is his “liberal principle of legitimacy”:⁵⁶

“Our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason” ([1993]2005: 137).

⁵⁶ Rawls ([1993]2005: 225; 1995: 175ff.) distinguishes between legitimacy and justice: although they draw on the same set of political values, they have different domains and legitimacy makes weaker demands than justice. Exercises of political power may be legitimate but unjust, but the converse is not possible. I focus on legitimacy in this thesis because it gives a standard for the acceptability of exercises of political power.

He adds that

“Only a political conception of justice that all citizens might be reasonably expected to endorse can serve as a basis of public reason and justification” (*ibid*).

In this subsection, I unpack the liberal principle of legitimacy in order to evince, in the next subsection, the criteria which a public reasons argument for the legal proscription of RSV must satisfy.

Rawls’ liberal principle of legitimacy is premised on a conception of the constituency of public reason as comprising *reasonable persons*, viz., persons who accept two main ideas. First, reasonable persons are, when among equals, “ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so” (Rawls [1993]2005: 49).⁵⁷ Second, reasonable persons accept the “burdens of judgment”, namely they accept the many “hazards involved in the correct (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life,” (*ibid*: 55–56) which explain how reasonable and rational people will permanently disagree about many matters of value and ethics.

Rawls believes that reasonable persons will accept what he calls the “idea of public reason”. The idea of public reason “belongs to a conception of a well ordered constitutional democratic society” (1997: 765). A well ordered society is the Rawlsian development of social contract doctrine. It is a society in which:

- (i) everyone willingly accepts and affirms the same principles of justice;
- (ii) the basic structure of society—namely, its basic political, social and economic institutions—is publicly known to satisfy these principles; and

⁵⁷ Note: the moral “duty of civility” also requires that citizens should explain to one another how, at least with regard to constitutional essentials and matters of basic justice, the political positions “they advocate and vote for can be supported by the political values of public reason” (Rawls [1993]2005: 217). Others object that the aims of public reason are best served by permitting debate and deliberation unconstrained by the duty of civility (e.g. Benhabib 2002: 108–112; Bohman 2003; Gaus 2009; Gaus and Vallier 2009; Vallier 2014, 2016). Nothing in my argument will hinge on the civility or otherwise of reasonable persons, so I remain neutral with respect to this feature of Rawls’ theory.

- (iii) citizens have a normally effective sense of justice, that is, they want to do what justice requires of them (Rawls 1999b: 4-5, §69; [1993]2005: 35).

The assumption of reasonable persons' "strict compliance" with principles of justice within a well-ordered society is part of what Rawls calls "ideal theory".

Under these conditions, in a public forum and concerning fundamental political questions, "[a] citizen engages in public reason [...] when he or she deliberates within a framework of what he or she sincerely regards as the most reasonable political conception of justice, a conception that expresses political values that others, as free and equal citizens might also be expected reasonably to endorse" (1997: 773).

The idea is that reasonable citizens realise they cannot reach agreement on the basis of their irreconcilable comprehensive doctrines. So when fundamental political questions are at stake, Rawls believes that reasonable citizens will appeal only to the kinds of "public reasons" which are acceptable to other reasonable citizens or, in other words, justifiable on the basis of political conceptions of justice.

§2.3.2. Political Conceptions of Justice

A political conception of justice has three characteristic features:

- (i) It applies to the "basic structure" of society, namely "a society's main political, social, and economic institutions, and how they fit together into one unified system of social cooperation from one generation to the next" (Rawls [1993]2005: 11).

Rawls takes the kind of society in question to be "a modern constitutional democracy" (*ibid*), which has an associated "criterion of reciprocity". This is why, as I discuss in the next chapter, Rawls' criterion of reciprocity acts as a constraint on any political conception of justice.

A political conception of justice has two features which demarcate political relationships in a constitutional democracy from other kinds of relationships. Firstly, because the basic structure is

presumptively a “structure of basic institutions we enter only by birth and exit only by death” (Rawls [1993]2005: 135-6), the political relationship is nonvoluntary, which distinguishes it from “associational” relationships which exist in, for example, religious organisations (*ibid.* 137). Secondly, in a constitutional regime “political power is ultimately the power of the public, that is, the power of free and equal citizens as a collective body” (*ibid.* 136). This, Rawls tells us, distinguishes political relationships from “the personal and the familial, which are affectional” relationships (*ibid.* 137). These features have sometimes been taken to denote a sharp public/private distinction in Rawlsian political liberalism (see, e.g., Okin 1994: 26). Although I challenge that assumption in the next chapter, one might naturally think, as have feminist critics of Rawls, that political conceptions of justice will have little to say concerning the appropriateness, from a legal perspective, of private sexual activities like RSV.

- (ii) A political conception of justice “is presented as a freestanding view” (*ibid.* 12), namely it is presented or expounded independently of any comprehensive doctrine.

The notion of presentation is key here: whilst Rawls assumes that “all citizens [will] affirm a comprehensive doctrine to which the political conception they accept is in some way related”, a political conception of justice “is presented as freestanding and expounded apart from, or without reference to, any such wider background” (Rawls [1993]2005: 12). This gives us another part of Rawls’ answer to the question of political liberalism, namely his account of stability as an overlapping consensus of reasonable comprehensive doctrines—that is, of comprehensive doctrines which “accept[...] a constitutional democratic regime and its companion idea of legitimate law” (Rawls 1997: 766). In an overlapping consensus, stability is achieved when the political conception of justice underlying the legal framework will be endorsed by all reasonable comprehensive doctrines such that the political conception “is a module [...] that fits into and can be supported by various reasonable comprehensive doctrines that endure in the society regulated by it” ([1993]2005: 12). The

arguments advanced in §2.2 for the legal proscription of RSV—depending fundamentally as they did upon Judeo-Christian sexual ethics or an associated account of natural law, the politics of love or a particular moral conception of good character—were, for example, *not* presented as freestanding from “background” doctrinal commitments.

- (iii) For a political conception of justice, its “content is expressed in terms of certain fundamental ideas seen as implicit in the public political culture of a democratic society” (*ibid.* 13).

This public political culture—comprising “the political institutions of a constitutional regime and the public traditions of their interpretation” (*ibid.* 13-14)—is distinguished from the “background culture” of civil society—comprising “comprehensive doctrines of all kinds” (*ibid.* 14). The arguments which I explored in §2.1, ones which supported the legal permissibility of RSV, appealed to ideas implicit in public political culture, such as the right to privacy. But the arguments explored in §2.2 for the legal proscription of RSV did not: rather, each argument appealed to ideas drawn from the background culture.

The considerations developed in this section illuminate a lacuna in the literature: no one, so far, has advanced an argument for the legal proscription of RSV which is acceptable by the standards of public reason. It is to the remedy of this omission that I turn in the next chapter.

3. Robots and Reciprocity

In the company of a living enigma man remains alone [...] This subjective game, which can go all the way from vice to mystical ecstasy, is for many a more attractive experience than an authentic relation with a human being.

de Beauvoir (1956: 263)

In this chapter, I articulate and defend an argument which is wholly new for the legal proscription of RSV: I contend that there is a public reasons justification for the legal proscription of RSV as part of the cultivation of those political virtues which are required by reciprocity in the just society. Recall from the last chapter that public reasons justifications, because they appeal to political conceptions of justice, must have three features: (i) they must establish that the matter is of concern for the basic structure of society; (ii) they must be presented independently of any comprehensive doctrine; and (iii) they must be expressed in terms of ideas which are implicit in the public political culture of a democratic society. Firstly, then, in §3.1, I explain that private sexual activities are, for Rawls, a matter of concern for the basic structure of society when engagement in them undermines reciprocity; RSV, I argue, would undermine reciprocity. Secondly, in §3.2, I show that my argument for the legal proscription of RSV may be presented independently of any comprehensive doctrine because RSV cultivates a *political* vice, that is, a character trait which is incompatible with citizenship in a just and stable constitutional regime. Thirdly, I highlight that the ideas underlying my argument for the legal proscription of RSV are implicit—and sometimes explicit—in both domestic and international law. §3.3 examines some problems with the implementation of this legal proscription through the criminal law, centring on the concern that the cultivation of political virtue seems to call

for something more than the criminalisation of political vice. To accommodate this criticism, I turn in the next chapter to advancing principles for the development of a legal framework for SARs.

My arguments in this chapter will draw upon the recent literature on epistemic injustice. There has been a great deal of interest in epistemology regarding the success conditions for testimonial exchanges, that is, what it takes for a hearer to come to know something on the basis of a speaker's say-so.⁵⁸ But it is well-known that testimonial exchanges sometimes go systematically wrong. This is because there are problematic patterns in who we tend to believe and in who we tend not to believe, that is, in the *credibility judgments* which underwrite testimonial exchanges. Epistemic injustice occurs when testimonial exchanges systematically go wrong due to facts about a person's social identity (e.g., their race, class or gender). The most widely theorised form of epistemic injustice is called "testimonial injustice".⁵⁹ Testimonial injustice occurs when someone's assertions are accorded less (or more) credence than they deserve because of prejudice of some kind, such as bias regarding their social identity (Fricker 2007: 7) or harmful "myths" which render certain explanations of events more likely than others (Medina 2012, 2013). Fricker (2003, 2007) develops an epistemic virtue ethics, discerning the virtue of testimonial justice and the opposing vice of testimonial injustice. The vice of testimonial injustice is a disposition to commit such acts of epistemic injustice. The virtue of corrective testimonial justice is a disposition to remain aware of, and compensate for, your prejudices by interfering with your estimation of the value of someone's testimony. This corrective virtue, Fricker (2003: 161) argues, is cultivated through social training. In

⁵⁸ Some argue that, at a minimum, I need to know what I'm saying (Audi 2006; Burge 1993, 1997; Schmitt 2006). But others reject this requirement (Lackey 2006, 2008; McKinnon 2013, 2015, forthcoming). On other views, I merely need to be reliable in saying true things (Goldberg 2010, 2015). But the dominant view appears to be that I need to be reliable *and* you have to have some reason to think that I'm reliable.

⁵⁹ The other main variety is "hermeneutical injustice", which occurs when one is wronged in their capacity as "a subject of social understanding", that is, when people are denied the conceptual and linguistic resources to make sense of and communicate their experience (Fricker 2007: 7). I focus on the perpetration of testimonial injustice but there is good reason to believe that rape myths also contribute to hermeneutical injustice against victims and survivors of sexual violence (Hänel 2020).

this chapter, I argue that, in some cases, epistemic injustice is a matter of concern for the basic structure of society and, in those cases, the virtues which characterise epistemically just testimonial exchanges are also political virtues which the state may legitimately seek to promote.

§3.1. The Personal and the Political

In this section, I argue that private sexual activities are a matter of concern for the basic structure of society when engagement in them undermines reciprocity (§3.1.1). One way to undermine reciprocity is to systematically undermine measures protecting the basic rights of some group of citizens. For example, in §3.1.2, I draw on the recent literature on pornography to argue that pornographic representations of sexual violence (PSV) systematically undermine reciprocity through undermining the ability of women and girls to advance important claims of justice. In §3.1.3, I argue that RSV would be a particularly pernicious kind of PSV and, as a result, would further undermine reciprocity.

§3.1.1. Justice Behind Closed Doors: A Question of Reciprocity

As we saw in the last chapter, the first hurdle faced by anyone advancing a case for the legal proscription of RSV consists in justifying the encroachment of political principles of justice into the private sphere of harmless sexual activities involving only consenting adults. Indeed, it has not always been apparent, especially not to feminist commentators on Rawls, that Rawls' political liberalism is friendly to the application of political principles of justice within the domestic sphere of human life at all. What hope, you might then think, is there for the application of political principles of justice to sex?

In order to address this concern, I respond to the common misconception that there is a sharp public/private divide implicit in Rawls' insistence that political conceptions of justice apply to the basic structure of society. As I explained in the last chapter, this misconception often derives

from the distinction which Rawls draws between those “political” relationships which characterise the basic structure and those “associational” and “affectional” relationships which characterise other institutions, including religious ones and the family. Okin, for example, concludes that “the distinction Rawls draws between the political and the nonpolitical coincides with his distinction between public and nonpublic” (1994: 27, 34; 2004: 1556). It would spell trouble for my argument if she were right since, as I explained in the last chapter, paradigmatically “nonpublic”—i.e., private—activities in constitutional democracies comprise those sexual activities which neither harm nor wrong other people, including, *a fortiori*, RSV. But the application of political principles of justice within the private sphere is, for Rawls, much less clear-cut than Okin suggested. On the contrary, political conceptions of justice impose substantive demands of justice on even the most private of human activities including, as I will show in the rest of this section, RSV.

In “The Idea of Public Reason Revisited”, Rawls sought to elucidate why political principles of justice constrain even those associational and affectional relationships which characterise the private sphere, arguing forcefully that “[i]f the so-called private sphere is alleged to be a space exempt from justice, then there is no such thing” (Rawls 1997: 791). This is because Rawls ascribes a certain normative priority to the political relation “of free and equal citizens [...] within the basic structure of society” (*ibid*: 769-770; [1993]2005: 12).⁶⁰ According to Rawls, whilst “the basic structure alone is the primary subject of justice, the principles of justice still put essential restrictions on the family and all other associations” (Rawls 1997: 791). Of course, RSV is not one of these affectional or associational relationships. But Rawls clarifies that *all* domains of human life are subject to the constraints of the political relation: these “principles defining the equal basic liberties and opportunities of citizens always hold in and through all so-called domains” (*ibid*). In particular, Rawls

⁶⁰ The priority of the political relation also seems to be conceptual: “[a] domain so-called, or a sphere of life, is not [...] something already given apart from political conceptions of justice [...] but rather is simply the result [...] of how the principles of political justice are applied” (Rawls 1997: 791).

was concerned to emphasise, “the equal rights of women [...] are inalienable and protect them *wherever they are*” (*ibid*, italics added). The normative priority of the political relation, then, means that there is no private sphere which is exempt from the most basic normative requirements of justice.

You might at this point respond that whilst the private sphere is not wholly exempt from the normative requirements of justice, the normative requirements which apply in our private lives are very basic. In particular, you might think that the normative requirements which are contained essentially in the political relation—independent of any particular political conception of justice—are insufficiently robust to justify anything so ambitious as a legal proscription on RSV. Granted, the normative requirements intrinsic in the political relation are basic. But, as I show in this chapter, even these basic requirements will suffice to establish my case.

The normative requirements of the political relation are those contained within the criterion of reciprocity, which acts as “the limiting feature” on all of the “many forms of public reason specified by a family of reasonable political conceptions” (Rawls 1997: 774). As Rawls explains, “the criterion of reciprocity [...] specif[ies] the nature of the political relation in a constitutional democratic regime as one of civic friendship” (*ibid*: 771).⁶¹ Reciprocity requires that, when citizens propose terms of cooperation, “those proposing them must also think it at least reasonable for others to accept them, as free and equal citizens, and not as dominated or manipulated, or under the pressure of an inferior political or social position” (*ibid*: 770). Because of this, Rawls believes that any reasonable political conception of justice must include three features:

R1. “a list of certain basic rights, liberties, and opportunities” (*ibid*: 774);

⁶¹ Rawls’ suggestive mention of civic friendship might seem to support the view that part of the moral basis for the idea of public reason is that the practice of public reason is a constitutive part of a valuable relationship (Ebels-Duggan 2010; Leland 2019; Leland and van Wietmarschen 2017; Lister 2013). However, there is reason to doubt that civic friendship alone can ground the idea of public reason (Billingham and Taylor 2023).

R2. “an assignment of special priority to those rights, liberties, and opportunities, especially with respect to the claims of the general good and perfectionist values” (*ibid*); and

R3. “measures ensuring for all citizens adequate all-purpose means to make effective use of their freedoms” (*ibid*).

This is because reciprocity imposes substantive demands upon the kinds of principles which free and equal citizens may reasonably be expected to accept when determining principles for a constitutional democracy: “[f]or what reasons can both satisfy the criterion of reciprocity and justify denying to some person religious liberty, holding others as slaves, imposing a property qualification on the right to vote, or denying the right of suffrage to women?” (*ibid*: 771).

It follows from my discussion so far that political principles of justice apply to private sexual activities when engagement in those activities would undermine the requirements which reciprocity imposes always and everywhere upon free and equal citizens. As I explained in the last chapter, RSV cases like Robotic Rape would not seem to imply the violation of any person’s basic rights, liberties or opportunities. Nor would RSV cases seem to undermine the special priority which reciprocity assigns to those basic rights, liberties or opportunities. My contention in this section is rather that RSV cases would undermine measures which aim to ensure adequate all-purpose means for women and girls to make effective use of their freedoms. In order to establish the truth of this contention, it will be instructive to see how existing forms of pornography undermine reciprocity in a similar way.

§3.1.2. Beyond the Porn Wars: Rape Myths and the Reporting Crisis

During the “porn wars” of the 1970s and 1980s, so-called “anti-porn” feminists sought to establish that pornography—namely, in broad terms, “sexually explicit material [...] primarily designed to produce sexual arousal in viewers” (West 2022)—either constitutes or causally contributes to the

perpetration of sexual violence against women and girls.⁶² Of course, one may undermine reciprocity through undermining or deprioritising the right of all persons to be secure from violence, including sexual violence (contravening R1 and R2). But I take no stand in this thesis as to whether the consumption of pornography undermines or deprioritises the basic rights of women and girls in this regard. Rather, I wish to develop a new line of argument concerning the relationship between pornography and sexual violence, one which focuses instead on how certain kinds of pornography undermine the ability of women and girls to successfully advance important claims of justice.

Besides undermining R1 or R2, another way to contravene the requirements of reciprocity is to undermine R3, that is, those measures which purport to ensure “adequate all-purpose means” for citizens to make “effective use” of their basic rights to be secure from sexual violence. These measures include the provision of adequate enforcement mechanisms to ensure that redress is effectively available to all persons in the eventuality that these basic rights are violated. To undermine reciprocity, then, it suffices to systematically undermine those enforcement mechanisms, rendering them wholly inadequate for their purpose. These enforcement mechanisms are famously already in a dire state when it comes to the enforcement of the rights of women and girls to be secure from sexual violence.⁶³ One problem is that crimes are often not reported.⁶⁴ The Women’s Aid Federation of England and Wales (WAFEW) (2021: ¶4) notes that one of the key obstacles to reporting is the existence of “pervasive rape myths within society which shape responses to survivors”. Another problem is that even when crimes are reported, the credibility of survivor testimony is often wrongfully diminished in court by appeal to rape myths. As WAFEW (2021: ¶12)

⁶² On equality and, relatedly, the elimination of sexual violence as a reason for pornography regulation: MacKinnon (1993).

⁶³ The UK Parliament Home Affairs Committee (2022: ¶43) reports that “victims and survivors of rape are being failed by our criminal justice system” due to “unacceptable” and “exceptionally low volumes of rape charges and prosecutions”.

⁶⁴ Fewer than one in six (16%) female victims and fewer than one in five (19%) male victims aged 16 to 59 years of sexual assault by rape or penetration since the age of 16 years reported it to the police (ONS 2021).

documents: “Myths and stereotypes are [...] perpetuated in the courtroom, and are inconsistently challenged by prosecutors, for example how a ‘true’ victim will behave in the aftermath of a rape”.⁶⁵ In addition, these rape myths interact with sexual history evidence concerning the victim which “attempts to imply to the courtroom that the survivor is [...] less credible as a witness” (*ibid*).⁶⁶ More generally, research shows that rape myth acceptance directly correlates with a tendency to dismiss allegations (Krahé, Temkin, and Bieneck 2007).

One motor for the propagation of rape myths in society is the consumption of pornographic depictions of sexual violence (PSV)—viz., pornographic representations of sexual violence for the user’s sexual gratification. It has been argued since the 1985 Meese Report that, among other things, widespread consumption of PSV in a society makes individuals more ready to accept rape myths and desensitises them to sexualised violence (Dwyer 1995: legal appendix; Einsiedel 1992: 266; Langton 1993: 310). Research in the psychology and sociology of sexual violence supports this claim: front-line workers working with sex offenders and survivors consistently document how widespread exposure to and consumption of PSV plays a significant part in shaping attitudes and behaviours towards women and girls which make sexual violence seem like a normal part of sex.⁶⁷ Recent research more generally shows that higher PSV incidence in a society changes sexual scripts to normalise sexualised violence towards women and girls. The theory of sexual scripts has its origins in Gagnon and Simon (1973) and, broadly, concerns the norms governing sexual behaviours in a social context. For example, research suggests that exposure to and consumption of PSV are associated with the normalisation of high-risk and violent sexual behaviours (Braithwaite *et al.* 2015); objectifying attitudes towards women (Sun *et al.* 2016; Fritz and Paul 2017); physical aggression

⁶⁵ See also Smith and Skinner (2012, 2017).

⁶⁶ See also McGlynn (2017).

⁶⁷ See, for example, Wyre (1992: 241-2), documenting psychological evaluation of sex offenders, and the recent report for the British Government Equalities Office which found that “All [...] Frontline workers [...] acknowledged [...] pornography as an influential factor for harmful sexual behaviours towards women and girls” in the sense that it “contributes to a conducive context facilitating these behaviours” (Upton *et al.* 2021).

(Shor and Seida 2019, 2020) and discriminatory attitudes, especially against Asian women (Zhou and Paul 2016). As a result, Vera-Gray *et al.* (2021: 1244) argue that “sexual violence is a normative sexual script in mainstream online pornography, with significant implications at a social level for understandings of the difference between sexual pleasure and sexual harm”.

It is instructive to analyse the problem with rape myths in terms of Haslanger’s distinction between the “manifest” concept of a term—that is, its explicit official and formal definition—and its “operative” concept—that is, the implicit definition that would be extrapolated from its actual usage in a given community (Haslanger 2005: 14). Rape myths, then, divorce the manifest concept (the legal definition of “rape”) from its operative concept (the actual usage of “rape” in a given community). As we’ve seen, this divorce contributes to the widespread and wrongful deflation of the credibility of women’s and girls’ testimony about sexual violence. This is a serious testimonial injustice. It undermines reciprocity because it undermines enforcement mechanisms designed to ensure “adequate all-purpose means” for citizens to make “effective use” of their basic rights to be secure from sexual violence. Indeed, the systematic failure to enforce the rights of women and girls to be secure from sexual violence means that women and girls as a social group face wholly inadequate means to make effective use of their basic freedoms. As Green puts it:

“Women who have not been, and may never be, victims of rape nonetheless modify their lives for fear of it, or because their parents fear it: they feel obliged to avoid certain places, to be careful going out at certain times, to give thought to how their dress and manner will be received—in general, to remain unwillingly vigilant. Few men have experiences that are remotely similar.” (2013: 491).

And rape myths, propagated at least in part by PSV consumption, are known to be part of the problem. What does all this have to do with the legal proscription of RSV?

§3.1.3. Autonomous Pornography

It is standard in the SARs literature to draw an analogy between pornography and SARs.⁶⁸ Indeed, by all accounts, SARs may be viewed as a new kind of pornography. It is easy to see why this is the case if pornography is, as mentioned above, “sexually explicit material [...] that is primarily designed to produce sexual arousal in viewers” (West 2022), “any object that has been manufactured to satisfy sexual desire through its sexual consumption or other sexual use” (Vadas 2005: 177) or, more emotively, “sex[ual activity] between people and things, human beings and pieces of paper, real men and unreal women” (MacKinnon 1993: 109).⁶⁹ Of course, the “material”, “object” or “thing” in question is traditionally assumed to be verbal, pictorial or videographic; but why not also robotic? This presumption would already give us good reason, based upon the arguments advanced in the previous section, to think that RSV would be a matter of concern for the basic structure of a just society.⁷⁰ But for reasons which I develop below, RSV would be an especially pernicious form of PSV from the perspective of rape myth propagation.

Naively, this might seem unlikely because of the ways in which SARs differ from existing kinds of pornography. As I explained in chapter one, SARs are social robots which, unlike existing pornography, look, speak and act rather like human sexual partners. The experience of enacting RSV might, then, be more lifelike. By themselves lifelike representations of sexual violence when only consenting adults are involved are not a matter of concern for the basic structure. As Gersen (2019: 1795) emphasises, “[i]t is hardly uncommon for human partners to engage together in various

⁶⁸ See, for example, Danaher (2017a, 2017b); Sparrow (2017); Eskens (2017); Gersen (2019).

⁶⁹ Elsewhere, MacKinnon describes pornography somewhat differently, as “the graphic sexually explicit subordination of women through pictures and words” (MacKinnon 1987: 176). In that case, we would need to expand the concept to include the subordination of women “through pictures and words and autonomous robots”. But this seems true to the spirit of MacKinnon’s account.

⁷⁰ This line of argument has not been developed in the literature but is suggestive of how a Rawlsian political liberal might justify the UK’s somewhat controversial criminalisation of possession of an “extreme pornographic image”, which includes some if not all depictions of rape (Criminal Justice and Immigration Act 2008: pt.5). For criticism of the Act: Carline (2011); Cowen (2015).

degrees of consensual role-play that includes fantasies of force, coercion, roughness, resistance, domination, and submission” (cf. Crépault and Couture 1980; Critelli and Bivona 2008). Unlike PSV, which propagates rape myths by promulgating ready-made sexual scripts and stereotypes through verbal, pictorial and videographic means, RSV might at first blush seem to offer more scope for human imagination, creating a more private experience in which the SAR’s user may experiment with different kinds of sexual fantasy. In other words, the *autonomous* nature of SARs might initially seem to open up new possibilities for those experiments in living which are already a legitimate part of human sexual relationships.⁷¹

In response, I wish to draw your attention back to the distinction which I drew in chapter one between general and narrow AI. If SARs were powered by general AI—that is, if they displayed the kind of sophistication, flexibility and generality in their intelligence which human beings do—I believe that the objection considered here might have some bite: in particular, if the autonomous nature of SARs were to involve the kind of creative or quirky autonomy which we see in adult human beings, then RSV might not be the kind of motor for propagating harmful stereotypes and rape myths that PSV seems to be.⁷² But SARs will not be powered by general AI for the foreseeable future. Rather, SARs will involve narrow AI endowed, as I explained in §1.2.3, with at least some bottom-up machine learning capabilities to enable them to respond like a human sexual partner to new and unexpected situations. And, the AI ethics literature reveals, nothing propagates harmful stereotypes and bias quite like bottom-up machine learning. One problem is that bottom-up machine learning bases the AI system’s future behaviour on regressions taken over potentially biased input data and therefore incorporates and augments existing biases (Forrest 2021; Kirkpatrick 2017;

⁷¹ For notes of optimism in this vein: Levy (2007).

⁷² SARs powered by general AI would raise other issues, which I touch on in chapter five.

Cohen ms).⁷³ As we've seen, the manifest and operative concepts of sexual violence are radically divorced in much of our sexual culture, a problem which is often colloquially referred to as "rape culture". It is therefore difficult to see how we could find sufficient unbiased training data to ensure that SARs would not learn—and learn with unprecedented efficiency—how to systematically reproduce rape culture in their sexual behaviours.⁷⁴ Another problem is that bottom-up machine learning is very difficult to audit because it inherently lacks transparency (Brennan and Oliver 2013: 558): some of the best machine learning techniques (known as "random forests" machine learning) result in wholly inscrutable "black-box" AIs, which threaten the insinuation of rape myths and harmful stereotypes (derived from the sexual scenarios of its training data) into the SAR's behaviour (Berk and Bleich 2013). And, unlike in the ready-made scenarios of existing pornography, the problematic biases ingrained in SARs are not even in principle scrutable to human users.⁷⁵

SARs, then, would be a new kind of autonomous pornography. Combined with their humanlike appearance and communication, the nature of their autonomy means that RSV would be a much more effective mechanism for propagating harmful stereotypes and rape myths than existing kinds of PSV. This mechanism undermines reciprocity in the just society and is, therefore, a legitimate concern for political principles of justice applied to the basic structure. This brings me to the second concern for the application of political principles of justice to the regulation of RSV: can arguments for the legal proscription of RSV be presented independently of any comprehensive doctrine?

⁷³ For rich discussion: Feldman *et al.* (2015). Cf. Sinnott-Armstrong and Skorburg (2021).

⁷⁴ There is a parallel to be drawn here between the effects of rape culture on *narrow* as compared to *general* AI and its effects on *minors* as compared to *adults*. Without an adequate level of autonomy, it appears that minors are much more susceptible than adults to uncritically reproducing sexual scripts and rape myths from mainstream pornography in their own sexual behaviours (Adarsh and Sahoo 2023).

⁷⁵ In limited cases, there may be ways to audit certain kinds of black-box AI for bias: see Adler and others (2016).

§3.2. Virtues of the Polity

In this section, I argue that, in addition to being a matter of legitimate concern for the basic structure, a legal proscription on RSV may be defended without appeal to any comprehensive doctrines and based upon ideas implicit in public political culture. First, I articulate and defend the Rawlsian distinction between moral virtues and vices and political ones (§3.2.1). Then, I show that in contexts where the perpetration of testimonial injustice undermines reciprocity, testimonial injustice there becomes a political vice (§3.2.2). §3.2.3 argues that legal measures cultivating the corrective political virtue of testimonial justice may be justified by appeal to ideas already implicit in public political cultures. As a result, the legal proscription of RSV is part of reasonable measures to promote reciprocity, which is a requirement of all political conceptions of justice.

§3.2.1. When is virtue political?

At this point in my argument, one might pertinently point out that lots of activities undermine (or sustain) reciprocity but do not thereby warrant legal proscription (or prescription). Your smoking cigarettes, for example, might set a bad example to others, who will thereby associate this habit with high intelligence and academic success; in this way, your smoking might undermine measures (such as anti-smoking campaigns) designed to ensure for all citizens “adequate all-purpose means to make effective use of their freedoms” (assuming that many freedoms cannot be so easily enjoyed by, for example, cancer patients). Nonetheless, it seems inappropriate for the state to ban smoking by appeal to the viciousness of setting a bad example. On the flip-side, daily Buddhist meditation practices for police officers might cultivate certain forms of compassionate attentiveness, which would sustain reciprocity by combating the testimonial injustices of rape myths. But even if this link could be empirically proven, it would still seem inappropriate for the state to legally prescribe daily Buddhist meditation practices by appeal to the virtue of compassionate attentiveness. So, the

objection goes, there are certain dispositions the cultivation or suppression of which is not an appropriate exercise of state power, irrespective of their effects on reciprocity.

This objection turns on a misunderstanding of the nature of public justification. The examples given above serve to highlight not the irrelevance of reciprocity in public justification, but rather that the reasons which we appeal to in public justifications matter too. It is true that it would be inappropriate for free and equal citizens to seek to justify exercises of political power by appeal to the viciousness of setting a bad example or the virtue of compassionate attentiveness. This is because the justification would be *presented* in terms of “moral virtues”, that is, those virtues which are appropriate “to the relations between individuals”, which hold in, for example, religious and familial circles (Rawls [1993]2005: 195). But it would *not* be inappropriate for free and equal citizens to seek to justify exercises of political power by appeal to the effects of certain virtues and vices on reciprocity. This is because the justification would be presented in terms of “political virtues”, that is, those virtues which are intrinsic to the political relation because they “characterise the ideal of a good citizen in a democratic state” (*ibid.* 194-5). The categories overlap: as Rawls explains, “political virtues are identified and justified by the need for certain qualities of character in the citizens of a just and stable constitutional regime” (*ibid.* 194 fn.29). But “[t]his is not to deny that these same characteristics, or similar ones, might not also be nonpolitical virtues insofar as they are valued for other reasons within various particular comprehensive views.” (*ibid.*). This is why, in the objection under consideration, the moral vice of setting a bad example and the moral virtue of compassionate attentiveness would also be matters of political vice and virtue provided that the empirics of their effects on reciprocity could be corroborated. This prompts the following question: can we justify the legal proscription of RSV by appeal only to political virtues or vices?

§3.2.2. Testimonial Justice as a Political Virtue

The disposition to commit testimonial injustice against sexual violence survivors, I have argued, undermines the political relation; it is, therefore, a political vice. The corrective political virtue is, as Fricker explains, the disposition to remain aware of and compensate for your prejudices, a virtue acquired through social training (Fricker 2003: 161). Now, the cultivation of political virtues is a legitimate exercise of political power, for Rawls, because that cultivation may be justified independently of comprehensive doctrines. As Rawls ([1993]2005: 195) explains:

“[I]f a constitutional regime takes certain steps to strengthen the [political] virtues of toleration and mutual trust, say by discouraging various kinds of religious and racial discrimination (in ways consistent with liberty of conscience and freedom of speech) it does not thereby become a perfectionist state. Rather, it is taking reasonable measures to strengthen the forms of thought and feeling that sustain fair social cooperation between its citizens regarded as free and equal”.

My first contention, then, is that a constitutional regime may by the same logic take certain steps to strengthen the political virtue of testimonial justice, by discouraging certain kinds of what may appropriately be called *gender* discrimination, including the testimonial injustices faced by sexual violence survivors. I discuss this notion of gender discrimination further in §3.2.3.

I argued in §3.1.3 that RSV would contribute to the development of the disposition to commit testimonial injustice against sexual violence survivors. So RSV would contribute to the development of a political vice. Steps to foster the corrective political virtue of testimonial justice, then, appear to be part of “taking reasonable measures to strengthen the forms of thought and feeling that sustain fair social cooperation between its citizens regarded as free and equal”. These steps, because their justification appeals to matters of concern for the basic structure and is

presented without appeal to any comprehensive doctrine, are not incompatible with political conceptions of justice. My second contention in this section, therefore, is that one such step appropriately involves the legal proscription of RSV. In order to assess this latter contention, it will help to explore to what extent the ideas underlying my proposal are implicit in public political culture.

§3.2.3. Conventions Concerning Discrimination

As Rawls explains, “[i]n a democratic society there is a tradition of democratic thought, the content of which is at least familiar and intelligible to the educated common sense of citizens generally” (Rawls [1993]2005: 14), which should be “seen as a fund of implicitly shared ideas and principles” which may be used in public reasons justifications. In order to show that the legal proscription of RSV which I defend in this chapter is indeed justified by appeal to public reasons, then, I need to show that its content is expressed in terms of certain fundamental ideas seen as implicit in the public political culture of a democratic society. In this subsection, I show that the idea of cultivating the political virtues associated with combating epistemic injustice are implicit within public political culture. Here, I put my case that the cultivation of those virtues is even a requirement under international law.

Within the public political culture of the UK, it is encouraging to see that the dangers which rape myths and stereotypes pose to the success of testimonial exchanges are recognised as a serious legal issue. The Crown Prosecution Service’s legal guidance for sexual offences devotes a whole chapter to “[t]ackling rape myths and stereotypes”, insisting that

“Rape myths and stereotypes [...] need to be identified and addressed where they arise in order to ensure a proper case-strategy and effective advocacy when presenting a case at trial” (CPS 2021: ch.4).

Moreover, the guidance recognises that

“technological advances” and associated “changes in sexual behaviours” in recent years mean that “there is an increased risk that myths and stereotypes may arise in rape cases in relation to the complainant, the offender, consent issues and/or the circumstances of the alleged offending” (*ibid*).

The guidance also cites *Miller v R* [2010], where the Court of Appeal endorsed the use of judicial directions to counter the risk of rape myths and stereotypes wrongfully deflating the credibility of survivor testimony. In other words, judges should direct counsel and juries to combat the testimonial injustices faced by sexual violence survivors, especially in light of how technological advances have affected sexual norms.

But at the global level, the call to cultivate political virtues to ensure adequate all-purpose access for women and girls to make effective use of their basic freedoms has been an explicit requirement under international law since the 1979 Convention on the Elimination of all forms of Discrimination Against Women (CEDAW), which has been ratified by 189 countries.⁷⁶ Indeed, CEDAW explicitly understands eliminating discrimination against women in terms of guaranteeing the “recognition, enjoyment or exercise by women [...] of human rights and fundamental freedoms in the political, economic, social, cultural, civil or any other field” (Art. 1). But CEDAW was and remains a radical document, calling on state parties to

“take all appropriate measures: [...] [t]o modify the social and cultural patterns of conduct of men and women, with a view to achieving the elimination of prejudices and customary and

⁷⁶ Note: the USA has merely signed it and is, therefore, not legally bound by CEDAW. Still, “[t]he signature [...] creates an obligation to refrain, in good faith, from acts that would defeat the object and the purpose of the treaty” (Vienna Convention on the Law of Treaties 1969: Arts. 10, 18).

all other practices which are based on the idea of the inferiority or the superiority of either of the sexes or on stereotyped roles for men and women” (CEDAW 1979: Art. 5).

Prejudice and stereotypes against women, CEDAW indicates, are not only personal vices but matters of global public political concern. Their elimination has therefore been a requirement under international law for several decades. One appropriate measure to achieve this end, which I have defended by appeal to public reasons acceptable to all reasonable citizens, is the legal proscription of RSV. How, though, should this legal proscription be implemented?

§3.3. Coercing Virtue

In this section, I explore certain practical and principled problems with the coercive enforcement of a legal proscription on RSV through the criminal law. I focus on the concern that the cultivation of political virtue seems to call for something more than the criminalisation of political vice. This latter concern prompts me to turn in the next chapter to advancing new principles for the development of a legal framework for SARs.

Firstly, the practical problems. As I discussed in the last chapter, Danaher (2017a) argues for the legal proscription and criminalisation of “rape and child sexual abuse robots”. Here, Danaher appears to be referring to sex dolls rather than SARs, because the robots in question are not autonomous. This is important because Danaher’s argument depends upon the assumption that the enactment of sexual violence with robots could in principle be prevented by banning the production, sale and use of all and only those robots which have been specifically designed for the purposes of enacting sexual violence (cf. Sparrow 2017). Privacy concerns might constrain the monitoring of individual doll usage, but Danaher believes that the production, sale and use of “rape and child sexual abuse robots” should be criminalised. What Danaher fails to consider, though, is what machine learning in “sex robots” will entail for their autonomous capacities. In particular, as I

explained in chapter one, we cannot legally proscribe RSV by legally proscribing the production, sale and use of SARs designed specifically for the purposes of enacting sexual violence. This is because even a tried-and-tested SAR which reliably consents enthusiastically to all sexual advances may withhold consent on any new occasion. How, then, can we hope to enforce the legal proscription on RSV?

Since any SAR is a possible site for RSV, one response might be to criminalise the production, sale and use of SARs altogether. I do not wish to rule out this proposal. But there are two reasons why I find it unsatisfying. Firstly, many believe that, at least as a general rule, something should be morally wrong in order to warrant criminalisation.⁷⁷ On this view, it might seem that criminalising the production, sale and use of SARs is overinclusive. Indeed, I have neither seen nor defended a compelling public reasons argument to think that the production, sale and use of SARs warrants criminalisation on its own terms; criminalising these activities, then, would be justified by appeal to the unenforceability of legal proscriptions on RSV specifically. Of course, some are happy to accept that the criminal law is often overinclusive. Classic examples of allegedly overinclusive crimes include statutory rape and drunk driving offences, the claim being that there are isolated cases of each offence which are not morally wrong (Green 2020). Because of this, Green argues that moral wrongness is not a precondition for criminalisation. But even if it is true that some crimes are overinclusive as Green (2020) suggests, this does not mean that we should embrace overinclusivity as a perennial feature of the criminal law. On the contrary, overcriminalisation would seem to undermine the message of condemnation which criminalisation intuitively conveys: as Coffee has it,

⁷⁷ This is what Duff (2014: 218) calls “negative legal moralism”. It is endorsed by, e.g., Duff (2014: 219); Feinberg (1984); Husak (2007: ch.2); Lamond (2007: 625). There is a well-known objection to negative legal moralism, which appeals to the distinction between two classes of crimes: ones which are *mala in se*—viz., “naturally evil as adjudged by the sense of a civilized community”, such as murder—and ones which are *mala prohibita*—viz., “wrong only because a statute makes it so”, such as speeding (*State v Horton* [1905]). But Duff (2018) believes that *mala prohibita* are compatible with the claim that conduct must be morally wrong in order to warrant criminalisation, even if that conduct need not be morally wrong prior to its legal proscription. For criticisms: Tadros (2016); Green (2020).

“stigma is a scarce resource” which diminishes in expressive power when overused (1991: 238). Because of this, I believe that we should be cautious about criminalizing the production, sale and use of SARs by appeal to the difficulty of enforcing the legal proscription on RSV more specifically.

Secondly, even bracketing concerns about overinclusion, whilst Rawls clearly saw the cultivation of political virtues as an appropriate exercise of political power, there is something odd about the coercive enforcement of political virtue through the criminal law. One way to think about this reflects on the ideal underlying political virtue in the first place, namely “the ideal of a good citizen in a democratic state” (Rawls [1993]2005: 194-5). It might seem to be in tension with this ideal if the democratic state were to coercively enforce the political virtues associated with good citizenship. More generally, the coercive enforcement of the virtues associated with good citizenship might seem to undermine the core idea behind public reasons arguments, namely the notion of respect for persons as free and equal citizens (Larmore 1999). Indeed, Rawls contemplates how the law may be conducive to—rather than coercive of—the development of certain political virtues, “strengthen[ing] the forms of thought and feeling that sustain fair social cooperation between its citizens regarded as free and equal” (Rawls [1993]2005: 195). It appears to me that he is gesturing here towards an alternative *educative* role which the law may play in the just society, one which guides without forcing the development of political virtues, remaining respectful of that special autonomy which—for now at least—accrues only to free and equal persons. I develop this idea in the next chapter.

§3.4. Concluding Remarks

In this chapter, I articulated and defended a public reasons basis for the legal proscription of RSV. This is because RSV would undermine reciprocity and is, therefore, a matter of concern for the basic structure of society (§3.1). In addition, the legal proscription of RSV may be presented

independently of any comprehensive doctrine and by appeal to ideas implicit in public political culture: RSV cultivates a particular political vice, which is condemned both domestically (in basic English legal institutions) and internationally (via CEDAW's widespread ratification) (§3.2). §3.3 explained why exclusive reliance on the criminal law when implementing the legal proscription on RSV might seem a little flat-footed in the cultivation of political virtue. In the next chapter, I explore another way to implement the legal proscription on RSV, one which attends to exactly what it means to be a legal person.

4. Robots and Persons

*If you prick us do we not bleed? If you tickle us do we not laugh? If you poison us do we not die? And if you wrong us
shall we not revenge?*⁷⁸

In the last chapter, I showed that there is a public reasons justification for the legal proscription of RSV. To complete my response to the RSVP, in this chapter, I sketch the contours of a legal framework for the implementation of that legal proscription. My contention is that the ascription of legal rights to SARs to be secure from RSV would be part of “taking reasonable measures to strengthen the forms of thought and feeling that sustain fair cooperation between [...] citizens regarded as free and equal” (Rawls [1993]2005: 195). An important corollary of this claim is that SARs are the kinds of entities which may appropriately bear legal rights. In other words, SARs should be ascribed legal personhood. In §4.1, I argue for a “functionalist” interpretation of legal personhood, one on which “a person is an agent who can perform effectively in the space of [legal] obligations” (List and Pettit 2011: 173).⁷⁹ According to this functionalist interpretation of legal personhood, I show in §4.2 that SARs should be ascribed legal personhood, including legal rights against RSV. I conclude in §4.3 that the ascription of legal personhood to at least some autonomous robots might be a legally efficacious and morally appropriate way to reflect our evolving relationships with them.

§4.1. The Grounds of Legal Personhood

It is often assumed that SARs would not be the kinds of entities which may appropriately bear legal rights because they do not possess the kinds of properties which are generally believed to underpin intrinsic moral status (cf., Darling 2014: 1 fn.4). Indeed, there is remarkable agreement that SARs

⁷⁸ These lines are spoken by Shylock, asserting his personhood, in Shakespeare’s *The Merchant of Venice* III(i).

⁷⁹ List and Pettit (2011) use the term “performative” instead of “functionalist”. But I prefer to leave the connection between the account of personhood and the philosophy of mind doctrine of functionalism explicit.

“do not meet the criteria for legal or moral personhood” (Danaher 2017b: 73) because they “are ‘mere’ machines” (Sterri and Earp 2021), that is, they are not “sentient being[s]” (Jecker 2020: 90), are “[in]capable of thinking and feeling” (Sparrow 2017: 467) or, as one Taoist analysis has it, have “no qi-energy” (Zhang 2021: 90). One exception to this concord appears in Frank and Nyholm (2017: 322). In their closing remarks, the authors gesture towards the possibility that SARs might be “brought into the legal community”, suggesting that we “take[...] seriously” the extension of “a form of ‘electronic personality’” to SARs. But no elaboration of this proposal is forthcoming in the SARs literature. My contribution in this chapter will be to offer principles to remedy this omission. To begin with, I suggest that legal personhood does necessarily depend upon the possession of specific intrinsic properties (§4.1.1). Rather, ascriptions of legal personhood should be based upon a constructive interpretation of law, one which both “fits” and “justifies” the extensional facts about actual ascriptions of legal personhood (§4.1.2). Functionalism about legal personhood satisfies these dual constraints (§4.1.3). In the next section, I apply functionalism to the question of whether SARs should be legal persons.

§4.1.1. Personhood Legalism and Personhood Moralism

According to one school of thought, which I call “personhood legalism” or “legalism” for short, the assumption in the SARs literature that legal personhood is a moral matter at all might appear utterly mistaken. For instance, in their defence of the legal personhood of certain kinds of AI systems, Chopra and White (2011: 155) suggest that “[l]egal entities are recognized as such in order to facilitate the working of the law in consonance with social realities” (cf. Turner 2019). Indeed, legalists take ascriptions of legal personhood to be derived only from other legal norms. The paradigm legalist is Kelsen, whose theory of legal personhood has three key aspects. First, being a

legal person is identical to having legal rights and obligations (Kelsen 1925: 63-4).⁸⁰ Second, since the law, for Kelsen, is an order of norms that prescribe the behaviour of natural persons (Kelsen 2005: 30-2; 1991: 28-9) and legal persons are simply bundles of legal norms, natural persons—whose behaviour is governed by legal norms—are legal persons.⁸¹ Third, Kelsen rejects any principled distinction between natural persons and artificial persons because the behaviour of so-called artificial persons, like corporations, encodes the behaviour of natural persons and is, therefore, the subject of legal norms. As a result, natural and artificial persons are both simply legal persons, that is, bundles of rights and obligations (Kelsen 2005: 174), which ultimately address the actions of individual human beings (*ibid.*: 171-178, 190-1).⁸² The refusal to ground legal principles in underlying moral principles is definitional of Kelsen’s “pure” theory of law, which “characterizes itself as a ‘pure’ theory of law because it aims at cognition focused on the law alone”, that is, the derivation of legal principles based only on other legal principles, independent of moral, political or scientific considerations (Kelsen 1992: 7; cf. Kelsen 2005). As a result, legalists of the Kelsenian tradition would be unmoved by arguments against ascribing legal personhood to SARs on the basis that SARs lack intrinsic moral status.

There appear to be two principal problems with legalism, one practical and the other more principled. First, the practical problem. There is an attractive principle of adjudication, famously developed in Dworkin’s *Law’s Empire*, on which “legal reasoning is an exercise in constructive interpretation” (Dworkin 1986: vii), that is, it “requires our judges [...] to treat our present system of public standards as expressing and respecting a coherent set of principles, and, to that end, to

⁸⁰ See also: Kelsen (1992: 46-7; 1945: 93-5; 2005: 172-3; 2015: 6); cf. Altwicker (2015: 236-9). Kelsen compares the separation of the legal person from its rights and obligations to the separation of a tree from its trunk, branches and blossoms: to do so would be impossible (2015: 172; cf. 1925: 62). As Kirste puts it, Kelsen explains the legal person “by making it disappear” (2000: 33).

⁸¹ Still, Kelsen grants that legal norms that address non-human beings are possible in theory. Indeed, he explores the practice of animal trials arising from animist beliefs and the possibility of training animals (Kelsen 1925: 62-3; 1923: 518-9; 1945: 3-4; 2005: 32-3; 1991: 89-91; 1992: 29).

⁸² On so-called “artificial” persons: Kelsen (1992: 48-9; 2005: 174-91).

interpret these standards to find implicit standards between and beneath the explicit ones” (*ibid.* 217).⁸³ But according to legalism there is no reason to treat our present system of public standards as expressing and respecting any coherent set of underlying principles. So this attractive adjudicative principle cannot get off the ground.⁸⁴ Because of this, Finnis critiques legalistic approaches on the basis that such approaches offer “no guidance” for “judges confronted by the practical problems of identifying law and rights after revolution” (Finnis 2011b: 26).⁸⁵ Whilst legalists might be content to bite this bullet, it is worth noting that many of law’s subjects want to see some kind of principled consistency in legislation and adjudication. Indeed, when the European Parliament tabled the legislative proposal to create a new legal status as “electronic persons” for certain autonomous robots, it was a central point of contention in the ensuing public debate that “[f]rom an ethical [...] perspective, creating a legal personality for a robot is inappropriate” (Open Letter to the EU Commission 2018). This leads nicely into the second, more principled problem: as Finnis argues, legalistic approaches seem to mistake “law’s point”, namely that ascriptions of legal personhood should translate “disciplined and critical concern with the interests and well-being of real people” (Finnis 2011b: 26).⁸⁶ How might legal personhood ascriptions reflect this kind of concern?

The school of thought concerning legal personhood which I call “personhood moralism”, or “moralism” for short, centres on the idea that, as a general rule, legal personhood depends upon moral personhood.⁸⁷ On the standard view, moralists assume that ascriptions of moral personhood—and, by extension, of *legal* personhood—depend upon an entity’s intrinsic properties,

⁸³ Dworkin (1986: 217) calls this the “adjudicative principle of integrity”.

⁸⁴ This is one way in which Dworkin’s adjudicative principle of integrity depends upon his “legislative principle of integrity”, which “asks those who create law by legislation to keep that law coherent in principle” (Dworkin 1986: 167). If the legislative principle is false, as it is according to legalism, it is unclear why judges should interpret law as morally coherent.

⁸⁵ In a similar vein, legalism quite closely resembles the “plain fact” view of law which Dworkin (1986: ch.1) critiques because it cannot explain or support adjudication in “hard cases” in law.

⁸⁶ This echoes Noonan’s (1976) classic case that judges who reasoned legalistically, viewing legal persons as “masks” rather than genuine “persons”, produced and reproduced the injustice of legal slavery for decades.

⁸⁷ For some apparent exceptions: *supra* fn.37.

such as dignity or sentience.⁸⁸ For instance, Finnis proposes a moralist conception of legal personhood according to which “all human beings [...] have the ‘right to recognition everywhere as a person before the law’ [UDHR: atcl.6]” (Finnis [2000]2013: 23) due to the intrinsic dignity of every human being.⁸⁹ As we saw above, moralism appears to be the working assumption in the SARs literature so far and the account of legal personhood which I defend in this section is a version of moralism.

But the standard assumption that moral personhood—and so legal personhood—is grounded in some intrinsic property like dignity or sentience does not fit with an important extensional fact about ascriptions of legal personhood, namely that courts across the Western world grant corporations legal personhood and have done so since the nineteenth century.⁹⁰ As a result, although a corporation lacks dignity or sentience, it can nonetheless enter into contracts, sue and be sued, own property and so on. Moralists tend to view the legal personhood of corporations as “simply another mode by which *individuals* or *natural persons* can enjoy their property and engage in business” (Hohfeld 1909: 197). Indeed, Finnis (2011b: 30) endorses this kind of position, arguing that, since “a corporation’s rights [...] are analysable without remainder into individuals’ rights”, we should understand the legal personhood of corporations as justified not by the corporation’s intrinsic properties *per se* but rather, ultimately, by appeal to the dignity of individual human persons. But the problem remains: corporations, on the standard moralist view, are *legal* but not *moral*

⁸⁸ For detailed discussion: Naffine (2009: 20). I bracket *rationalism*—viz., the view that legal personhood is grounded in an entity’s rationality—since this ostensibly under-extends, ruling out those natural persons who do not meet some standard of rationality, including the severely mentally ill, infants, those suffering neurological degeneration or comatose adults.

⁸⁹ Although Naffine calls this kind of position “religionism, as the Universal Declaration of Human Rights makes clear, considerations of intrinsic dignity need not be grounded in any world religion.

⁹⁰ See, e.g., von Savigny (1840: §60, §85) (Germany); *Trustees of Dartmouth College v Woodward* [1819], *Santa Clara Co v. Southern Pacific Railroad* [1886] (U.S.A); Blackstone (1876: 455, 186-7) (England). For an overview of the historical corporate personality debate: Conard (1976: 417-45). Cf. Freund ([1897]2000); Dewey (1926); Machen (1911); Hallis (1930). Olivecrona (1928) provides a helpful summary of “classical” Continental theories of the nineteenth century.

persons, which is in tension with the core claim of moralism.⁹¹ Such considerations are not decisive, but they suggest that dismissing the legal personhood of SARs based upon their lack of certain intrinsic features is undermotivated. As the case of corporations shows, legal personhood need not depend upon those intrinsic properties. On what basis, then, should we assess the proposal to ascribe legal personhood to SARs?

§4.1.2. Legislative Integrity

In *Law's Empire*, Dworkin proposes what he calls the “legislative principle of integrity”, which “asks those who create law by legislation to keep that law coherent in principle” (Dworkin 1986: 167), that is, “to try to make the total set of laws morally coherent” (*ibid.*: 176). On this view, there are two dimensions on which we should assess the legitimacy of any new legal proposition: firstly, does the proposition *fit*—i.e., is it *consistent*—with the norms conveyed by legal institutions? Secondly, is the proposition *justified* by (a) sound moral principle(s)? In light of the legislative principle of integrity, then, I want to reflect, in the context of SARs, on the European Parliament’s proposal that at least some autonomous robots should be granted legal personhood. Integrity calls us to evaluate this proposal along two dimensions:

Q1. Would the ascription of legal personhood to SARs fit with the norms conveyed by legal institutions?

Q2. Would the ascription of legal personhood to SARs be justified by a sound moral principle?

To begin with, then, we need some account of the norms conveyed by legal institutions concerning ascriptions of legal personhood.

⁹¹ Finnis might respond that corporations are not central cases of legal personhood but merely peripheral ones. But, as I show, the legal personhood of corporations seems to be a basic extensional fact about personhood in contemporary legal systems, one not to be confined to the conceptual peripheries.

Because my thesis is primarily philosophical rather than legal, I will not offer a complete answer to Q1. Instead, I restrict myself to the stipulation of just two familiar extensional facts concerning ascriptions of legal personhood which any account of the legal personhood of SARs must fit and justify. I focus on extensional facts about legal personhood ascriptions in contemporary Western legal systems, including common law (the UK and US) and civil law (Germany, France, Italy, Spain, Austria) jurisdictions. This makes sense for reasons of space and the contingent truth that, as we have seen, contemporary political discussions of AI legal personhood concern primarily these kinds of systems. Two extensional facts about legal personhood in these legal systems are as follows:⁹²

F1. All human beings who have been born, are currently alive and are sentient are legal persons. This is because: only human beings are so-called “natural persons”;⁹³ being born alive is generally sufficient if not necessary to warrant the ascription of legal personhood;⁹⁴ dead people are not legal persons;⁹⁵ it is uncontroversial that born, living, sentient humans are legal persons.⁹⁶

F2. In addition to natural persons, there are so-called “artificial persons”, such as corporations. In general, there appear to be three principal ways in which the legal personhood of corporations has historically been understood (Orts 2013):

- (i) *Concession (or fiction) theories*: corporations exist only as creations of the state or the sovereign.
- (ii) *Participant theories*: corporations have legal rights and/or legal duties which encode the legal rights and/or duties of the individual legal persons who compose the corporation.⁹⁷

⁹² For elaboration of these and other extensional facts in contemporary Western legal systems: Kurki (2019: *Introduction*).

⁹³ See: Berg (2007: 373); cf. German Civil Code (§1); Austrian Civil Code (§16).

⁹⁴ Being born is explicitly necessary for natural personhood in French, German and Italian law. See, e.g., German Civil Code §1; Italian Civil Code Pt.1:§1. In Spanish law, the neonate must also live for 24 hours outside of the womb (de las Heras Ballell 2010). An American landmark case, defining fetuses as non-persons under the 14th Amendment of the Constitution was *Roe v Wade* [1972].

⁹⁵ Although Kramer (2001) suggests that dead people might possess some limited rights. Cf. Stemplowska (2020).

⁹⁶ There is controversy regarding the legal personhood of anencephalic infants since they lack sentience. Anencephalic infants are legal persons in the US (Berg 2007: 377-379) and Germany (Neuner 2013: 651).

- (iii) *Real entity theories*: a corporation is a real entity over-and-above its members and has legal rights and/or duties in a given domain in virtue of the role that it plays in that domain.

The classic judgment cited as laying out the concession theory of the corporation is *Trustees of Dartmouth College v Woodward* [1819], which states that corporations exist only in the “contemplation of law” (636). More generally, as Kurki (2019: ch.1) shows, concession theories were popular until the middle of the nineteenth century (Blackstone 1876: 186-7; von Savigny 1840: §85).⁹⁸ But concession theories were largely supplanted by participant theories in the late nineteenth century. For example, in *Santa Clara Co v. Southern Pacific Railroad* [1886], the Court is clearly “looking through the corporation to its shareholders and their constitutional rights” (Finnis 2013: 46, fn.4).⁹⁹ The last century has seen a decisive and significant deepening of the legal status of corporations, endowing corporations with a status that seems to me best described by real entity theories. This interpretation of legal personhood has, for example, been the dominant one in the US Supreme Court in recent years, which “attribute[s] constitutional rights to corporations precisely as real entities in themselves” (*ibid*; cf. Horwitz 1985: 178, 192, 189). Indeed, the Supreme Court has ruled that business corporations have the same free speech rights as natural persons in election campaigns (*Citizens United v. Federal Election Commission* [2010]); religious freedom under a federal law not to provide birth control coverage in employee health plans (*Burwell v. Hobby Lobby Stores, Inc.* [2014]) and, arguably, certain religious freedom rights under the First Amendment (*Masterpiece Cakeshop, Ltd. v. Colorado*

⁹⁷ Cf. Jones (1999) on the “collective conception” of groups.

⁹⁸ Concession theories have an even longer pedigree: in 1246, Pope Innocent IV declared that corporations were *persona ficta* (List and Pettit 2011: 170).

⁹⁹ See also Kelsen’s theory of corporate legal personhood, on which the legal personhood of corporations is no more a legal fiction than is any other kind of legal personhood and is entirely explainable in terms of its role in regulating human conduct in law.

Civil Rights Commission [2018]). As a result, I understand the legal personhood of corporations signalled in F2 according to a real entity theoretic approach.¹⁰⁰

Based upon my arguments in §4.1.1, it appears that legalism *fits* but does not *justify* F1 and F2. Meanwhile, the standard interpretation of moralism can *justify* F1 and F2, but does not *fit* them both. With these considerations in mind, how should we interpret F1 and F2?

§4.1.3. Personhood Functionalism

One way to amend the moralist account of legal personhood in order to better accommodate F2 would be to ground moral personhood in something other than dignity or sentience. This is what List and Pettit do in their functionalist account of moral personhood as *moral agency*: “a [moral] person is an agent who can perform effectively in the space of moral obligations” (List and Pettit 2011: 173). As I unpack this definition, I explain why certain groups, which I call corporations, satisfy it.

Firstly, an agent is defined in the following way. For some entity X and some context C, X is an agent in C iff (*ibid.* 19-20):

- A1. X has representational states that depict how things are in the environment in C;
- A2. X has motivational states that specify how X requires things to be in that environment; *and*
- A3. X has the capacity to process its representational and motivational states, leading it to intervene suitably in the environment in C whenever that environment fails to match a motivating specification.

List and Pettit argue that different organisational structures in groups of people mean that certain groups—which I denote by the term “corporations”—are agents in certain contexts. First,

¹⁰⁰ Considerations advanced in the next subsections in favour of a functionalist conception of legal personhood offer further support for the conceptualisation of corporations as real entities over and above their individual members.

corporations have mechanisms for forming representational states. For example, some corporation Y's public relations team may collect polling data which informs the judgment ascribed to the corporation as a whole that its latest publicity campaign is transphobic. Second, corporations have motivational states, that is, preferences about how the group requires the state of affairs to be. Y's constitution, for instance, might require that all publicity campaigns comply with an ethos of inclusivity. Third, corporations have decision-making mechanisms which endow the group with the capacity to intervene suitably in a given situation on the basis of its judgments and preferences about that situation: Y might discontinue the publicity campaign in question on the basis that its executive committee has voted against it.

Secondly, to perform effectively in the space of moral obligations, an agent must satisfy a further three conditions, which are necessary and sufficient for the agent to “contract[...] obligations to others and [...] derive[...] entitlements from the reciprocal obligations of others” (List and Pettit 2011: 173). For some agent X in C, X is a moral person in C iff (*ibid.* 155, ch.8):

A4. X faces a normatively significant choice in C;

A5. X has the understanding and access to evidence required for making judgments about the options in C;

A6. X has the control required to make decisions in C.

For example, Y's accounting team might inform Y's judgment that the transphobic publicity campaign is exceptionally profitable and Y's constitution contains a clause allowing that, when they are exceptionally profitable, the discontinuation of publicity campaigns non-compliant with an ethos of inclusivity is a matter for the discretion of Y's executive board. Y therefore faces a normatively significant choice between retaining a transphobic, profitable campaign or discontinuing the campaign in favour of an alternative non-transphobic one. Y's policy analysis team informs Y's

judgment that (1) the current campaign is exceptionally profitable *because* it is transphobic and (2) it is therefore overwhelmingly likely that an alternative campaign would be less profitable but (3) there is good evidence to suggest that corporations which consistently display inclusive values in their public relations out-compete others over ten-year time frames. Y's executive board therefore votes on the options and on this basis Y decides to continue the campaign. This illustrates one way in which a group agent performs effectively in the space of *moral* obligations.

The functionalist account of legal personhood which I propose grounds the legal personhood of some entity in its moral agency, understood according to the functionalist approach glossed above. In other words, a legal person is a moral agent which is able to perform effectively in the space of *legal* obligations. To complete the illustration of my proposal, a claimant, Z, could meaningfully bring a legal case against Y, claiming that the campaign violates their legally protected right in that jurisdiction to be protected from discrimination on the basis of their gender identity. This right correlates to a duty borne by Y not to engage in certain forms of speech or expression. Y's defence might hinge upon its legally protected right, as a legal person, to freedom of speech or expression. The court—another kind of corporation—might then need to decide whether to grant Z's right or Y's right precedence in this context. Such considerations are of course not decisive, but I believe they give us sufficient reason to grant that the functionalist definition of legal personhood fits F2.

One might worry, however, that the functionalist interpretation does not fit F1 since not all natural persons perform effectively in the space of legal obligations. Think, in particular, of those who only bear what is called “passive” legal personhood, for instance infants (Kurki 2019: 94-96; MacCormick [1988]2007: 86). To resolve this concern, I will gloss what “passive” and “active” legal personhood standardly include. The Hohfeldian incidents paradigmatically associated with passive

legal personhood include certain fundamental claim-rights (protections of life, liberty and bodily integrity), the capacity to own property, victim status in criminal law and legal standing to enforce one's claim-rights in court through a representative. The Hohfeldian incidents paradigmatically associated with "active" legal personhood—a legal status which typically accrues to human adults of sound mind—include criminal-law and tort-law responsibility and the capacity to enter into contracts (Kurki 2019: 94-96; MacCormick [1988]2007: 86, 93). But it appears to me that the incidents associated with passive legal personhood are sufficient for an entity to perform effectively in the space of legal obligations (albeit primarily as rights-holders rather than duty-bearers). So the functionalist interpretation of legal personhood seems to fit F1, too.

Another objection, however, is that functionalism achieves fit with F1 and F2 only at the cost of emptying the concept of legal personhood of all its normative content. Indeed, one might ask in what sense the functionalist account of legal personhood actually *justifies* any ascriptions of legal personhood since it does not depend upon any intrinsic, *morally relevant* features of an entity, such as dignity or sentience. My response has two parts. Firstly, a note of concession. To some extent, it is true that the functionalist account of legal personhood does not depend upon the entity to which it is ascribed having any *intrinsic* moral status, that is, moral status which an entity has independently of its relationships with other entities. To illustrate, a corporation has no intrinsic moral status but a human baby does: destroying the former, for example, seems to be wrong only in virtue of the effects which that destruction has on entities like natural persons who may lose their jobs and suffer; destroying the latter, by contrast, seems to be wrong in itself. But secondly—and this is crucial—my account *does* depend upon the entity to which it is ascribed having at least *extrinsic* moral status, that is, moral status which an entity has in virtue of its relationships with other entities. In the former example, the corporation seems to have extrinsic moral status in virtue of its relationships with various intrinsically morally valuable humans. The functionalist conception of

legal personhood has not emptied legal personhood of its normative content; on the contrary, functionalism promises an account of legal personhood which is sensitive to the normative significance of our relationships with non-human entities. In order to further develop this latter point, I wish to clarify the normative content of the functionalist conception of legal personhood through applying the theory developed in this section to the case of SARs.

§4.2. Oh Brave New World: Towards a Legal Framework for SARs

In the previous section, I articulated and defended a functionalist conception of legal personhood, one which best satisfies the dimensions of *fit* and *soundness* on which the ideal of integrity evaluates the development of law. In this section, I show that SARs should be legal persons according to this functionalist conception (§4.2.1). In §4.2.2, I distinguish between non-derivative and derivative legal rights—that is, respectively, legal rights which are grounded in the rights-holders own interests and those which are grounded in the interests of others—and argue for the ascription of derivative rights against RSV to SARs. §4.2.3 defends my response to the RSVP against certain objections.

§4.2.1. The Legal Personhood of SARs

To begin with, I argue that SARs are *agents*, that is, they satisfy (A1)-(A3).¹⁰¹ There appear to be two principal ways to justify this claim. Firstly, by drawing an analogy with the case of corporations. To see how this works, consider Dennett's (1987, 2009) "intentional stance" test for agency, according to which "[a]nything that is usefully and voluminously predictable from the intentional stance is, by definition, an intentional system" (Dennett 2009: 339). Theorists of group agency have argued that passing the intentional stance test is good evidence that a system satisfies (A1)-(A3).¹⁰² For instance,

¹⁰¹ Indeed, the illustrative example of agency in List and Pettit (2011) is a much simpler robot than a SAR. List and Pettit ask us to consider a "small robotic device" which moves around on a table-top covered in cylinders, some upright, some on their sides. Suppose that the robot "scans the area with its bug-like eyes, then moves about and, using its prosthetic limbs, puts any cylinders lying on their side into an upright position" (2011: 19). List and Pettit claim that the robot satisfies conditions (A1)-(A3) and is, therefore, an intentional agent in this context.

¹⁰² For arguments to this effect, see: Tollefsen (2002, 2015) and List and Pettit (2011).

the kinds of entities which we count as artificial persons in law (e.g. firms, corporations and other organised collectives) are “usefully and voluminously predictable from the intentional stance”. In addition to this, List (2021) shows that any AI system which passes a suitably specified version of Turing’s “imitation game” test will satisfy Dennett’s intentional stance test. Specifically, Turing proposed that when assessing whether a system is “intelligent”, we should replace the question “Can machines think?” with the more tractable question “Are there imaginable digital computers which would do well in the imitation game?”, where this means displaying behaviour that is indistinguishable, in an observer’s eyes, from that of an intelligent agent (Turing 1950: 442). Different specifications of the imitation game correspond to different versions of this test. In Turing’s own version, the system must imitate the responses of a human interlocutor in written communication via an instant-messaging system. Chat GPT seems to satisfy this version of the test. Because of the more intimate context of their use, the appropriate version of the Turing test for SARs would require more sophisticated capabilities. But, as I explained in chapter one, whilst autonomous robots which adequately mimic the behaviour of a real human sexual partner do not yet exist, the requisite technology and corporate interest in their development seem to. As a result, I submit that SARs would satisfy a suitably specified imitation game which, as List points out, offers good reason to infer that they would pass Dennett’s intentional stance test and, as a result, satisfy (A1)-(A3).

Secondly, List (2021) argues that group agents like corporations are not just *analogous* to but rather *special cases* of AI systems, where the “hardware” supporting their artificial intelligence is social rather than electronic. This draws on Parikh’s (2002) concept of “social software”, which captures the idea that certain social phenomena are appropriately conceptualised as computational processes implemented in social systems. Parikh uses this term to refer to a broad range of social phenomena, including institutional decision-making and voting mechanisms, but he does not focus on group

agency. List extends the notion of social software, arguing that “group agents are socially implemented AI systems” (List 2021: 1221). Indeed, as functionalist philosophers of mind have long noted, agency is a multiply realisable phenomenon. In this vein, List contends that agency admits at least three different kinds of “hardware systems”:

- Biological systems (e.g. in human and non-human animals which display agency);
- Electronic systems (e.g. in AI systems such as autonomous robots);
- Social systems (e.g. in the case of group agents).

List notes that the boundaries between these different kinds of hardware may become increasingly blurred with the arrival of controversial new technologies such as biological robotics, human enhancement and technologically augmented social systems. But such considerations need not detain us here since I believe that my arguments so far suffice to show that SARs, like group agents, would satisfy (A1)-(A3).

In addition to this, SARs appear to be *moral* agents, that is, agents which perform effectively in the space of moral obligations. To illustrate, consider a SAR in a sexual context. Because I have argued that we may take the intentional stance towards the robot, I will use phrases such as “the SAR judges/intends/wants/acts”. Suppose that the SAR, in this context, judges that its human user is aroused, aggressive and attempting to initiate a sexual encounter on the basis of the user’s body language, speech and facial expressions. Since it is autonomous, the SAR faces the choice as to how to respond: should it consent enthusiastically, acquiesce, say no or shut down? Any one of these choices, as I showed in the last chapter, is normatively (if not yet legally) significant. As a result, the SAR satisfies (A4). The SAR understands that its user’s behaviour is aggressive and has sufficient evidence from past sexual encounters to judge that saying no or shutting down increases the risk of damage to its systems, whilst signalling acquiescence or consent encourages the same kind of future

behaviour in the user. As such, the SAR seems to satisfy (A5). Based upon its internal evaluation of what would minimise harm to its systems in the long term, the SAR determines its response, perhaps saying “no” to its user’s sexual advances. This seems to satisfy (A6). I submit that SARs, then, seem to be moral agents according to the functionalist conception of moral agency. According to the functionalist interpretation of legal personhood, which I defended in the last section, SARs may also, therefore, be meaningfully ascribed legal personhood. In other words, SARs would be appropriate holders of legal rights or bearers of legal duties. In the next subsection, I turn to the consideration of which kinds of legal rights SARs may appropriately bear; I defer consideration of the possibility that SARs might also bear certain kinds of legal responsibilities until the next chapter.

§4.2.2. Robot Rights

As I explained in the last section, according to the functionalist conception which I have defended, the ascription of legal personhood to some entity does not depend upon the intrinsic moral status or value of that entity. Entities, such as corporations, which lack intrinsic moral status may therefore be legal persons in virtue only of their extrinsic moral status. The distinction between intrinsic and extrinsic moral status sustains a second distinction between non-derivative and derivative legal rights. Non-derivative legal rights are those which a legal person has in virtue of her intrinsic moral status. For example, the kinds of fundamental protections which all natural persons have for their lives, liberty and property are non-derivative rights. Non-derivative rights are typically justified by appeal to something like the moral value of certain interests of the rights-holder: X’s non-derivative rights protect those interests of X which are important enough to generate legal duties on others to ensure their protection. Derivative legal rights, on the other hand, are those which a legal person has in virtue of the kinds of values or goods that they promote. For example, the property rights of corporations are typically justified by appeal to the desideratum of an efficient system of economic

exchange (Turner 2019). In this subsection, I argue for the ascription of derivative legal rights to SARs against subjection to RSV.

To begin with, I assume that SARs do not possess the kinds of intrinsic features which are generally believed to underpin intrinsic moral status, features such as dignity or sentience.¹⁰³ This is because, as I explained in chapter one, SARs involve only narrow AI and not that general AI which might be, in some sense, “conscious” (Danaher 2017a; Eskens 2017). As a result, any legal rights which we ascribe to SARs will be *derivative*, that is, justified by appeal to the values and goods which those rights promote. The central valuable good on which I want to focus is that which I showed in the last chapter was a central matter of concern for the basic structure of society, namely reciprocity. I argued that RSV undermines reciprocity because it undermines measures designed to ensure adequate all-purpose means for all citizens to make effective use of their freedoms. Because of this, I argued that reciprocity depends upon the elimination of RSV. The legal proscription of RSV, as a result, was shown to be part of reasonable measures for the cultivation of political virtue. One way to do this, I propose, is to create a legal right of SARs to be protected from subjection to RSV, a derivative right justified by appeal to its efficacy in protecting the valuable good of reciprocity. The question, then, becomes whether the ascription of such a legal right to SARs would, indeed, be an efficacious way to achieve this valuable social end.

There are several ways to justify the view that the ascription of this legal right to SARs would support reciprocity. In general, one might appeal to the positivist assertion that law, insofar as it claims to be law, asserts authority to give its subjects content-independent reasons for action (see, e.g., Raz 1994). As a result, there is something intrinsically action-guiding—albeit not necessarily *legitimately* so—about the ascription of a legal right to SARs proscribing some form of conduct

¹⁰³ In the next chapter, I relax this assumption to briefly explore the kinds of normative issues which would be raised by the more futuristic possibility of artificial general intelligence.

against it. Burge-Hendrix (2007) draws attention to a certain “educative” function of the law implicit in this positivist thesis, according to which part of the law’s function is to “make its subjects aware of the set of norms that structure[...] social and political life, and to persuade them to follow those norms” (243). Indeed, in his discussion of how social norms may solve collective action problems, Sunstein (1996: 2031) contends that the law may be used “as an effort to produce adequate social norms”, even when unaccompanied by much in the way of enforcement. He takes as an example laws which forbid littering and require that people clean up after their dogs. Although such laws are rarely enforced through the criminal law,

“they have an important effect in signalling appropriate behavior and in inculcating the expectation of social opprobrium and, hence, shame in those who deviate from the announced norm” (*ibid.* 2032).

As a result, with or without enforcement, laws may “help to reconstruct norms and the social meaning of action” (*ibid.*). In a similar way, although there might be practical or principled difficulties with enforcing the legal proscription on RSV in criminal law, as I discussed in the last chapter, ascribing a legal right to SARs to be protected from RSV may help to reconstruct sexual norms and the social meaning of certain violent sexual behaviours in ways which would protect the basic rights of women and girls. This kind of thinking, moreover, underlies Green’s (2013) case that law may be used to shape social norms and, in particular, to help foster better sexual norms. But he writes:

“thinking that *only* rape law stands a chance of shaping attitudes to rape is no better than jurisprudential homeopathy. In fact, it may be that the most influential laws in shaping our social morality here will not be the laws of rape at all [...] We should not assume that only like cures like: a wide variety of legislative initiatives can shore up the background conditions we are concerned about” (Green 2013: 492).

In other words, we are called to be not only open-minded in our consideration of which legal measures might help to foster positive social change; we are also asked to be thorough-going in our commitment to ensuring that legislative initiatives in all areas of life, including the most intimate ones, support the development of social norms conducive to reciprocity. My proposal is that the ascription of legal rights to SARs proscribing RSV is supportive of this valuable aim.

§4.2.3. Objections and Replies

One objection to my argument for the ascription to SARs of a legal right against RSV appeals to the standard, interest theory of legal rights. The most oft-cited version of this theory is Raz's:

“‘X has a right’ if and only if X can have rights, and, other things being equal, an aspect of X's wellbeing (his interest) is a sufficient reason for holding some other person(s) to be under a duty” (Raz 1986: 166; 1994: 266–269).

Raz continues: “[a]n individual is capable of having rights if and only if either his wellbeing is of ultimate value or he is an ‘artificial person’ (e.g. a corporation).” I have argued that autonomous robots like SARs should be artificial persons. But this leaves two questions. First: in what sense could a SAR be said to have “wellbeing” or “interests”? Of course, if *consciousness* or *sentience* is required for an entity to have wellbeing or interests, then the ascription of legal rights to SARs, according to the Razian interest theory, is dead in the water. But as I explain below, Raz does not believe that wellbeing or interests—at least for *artificial* persons—depend upon these kinds of intrinsic features. Second, then, even if the first question may be answered, how could a *robot's* interest be “a sufficient reason” for holding *humans* to be under a corresponding duty?

To answer the first question, I draw again on the analogy between SARs and artificial persons such as corporations. The sense in which *corporations* have interests on the Razian account is intrinsically tied up with their legal rights and duties:

“Corporations also have interests determined by their purposes, powers, and duties. It is true that protecting these interests is not intrinsically valuable. Nevertheless, corporations [...] have rights in the same sense as other individuals. They have rights if and only if their interests are sufficient to justify holding others to be subject to duties” (Raz 1995: 259).

One interpretation of what Raz says here is to define corporate rights in terms of corporate “purposes, powers, and duties”. But this would be odd because it seems to misconstrue the justificatory story of corporate rights. As Pallikkathayil (2016: 149) explains: “artificial persons [...] are created as packages of powers, rights, and duties, with each of these tailored to fit the others” such that “[e]xplaining the rights in terms of the duties [...] miss[es] the way in which the rights and duties are justified altogether (or not at all)”.¹⁰⁴ This kind of interpretation of Raz captures the interdependence of legal rights and duties according to the functionalist conception of legal personhood: that corporations, like SARs, perform effectively in the space of legal obligations—that is, they are legal persons—is inseparable from the ascription to them of certain legal rights and duties. But the *interests* of artificial persons like corporations and, I suggest, SARs in achieving certain ends do not seem, on the Razian account, to be explicable independently of their legal rights and duties. This prompts the next question, which concerns the value of interests like these.

To answer the second question, I will briefly explain how certain kinds of *instrumentally* valuable interests which accrue to legal persons on Raz’s account may be sufficient to ground a duty. As Raz explains, “[r]ights protect the interests of right-holders. But these interests need not be intrinsically valuable. The reason for protecting them may be that by doing so one does protect the interests of others” (Raz 1995: 259). He elaborates on this idea of instrumentally valuable interests as follows:

¹⁰⁴ I return in the next chapter to the question of what legal duties SARs might bear.

“The rights of journalists [...] to protect their sources are normally justified by the interests of journalists in being able to collect information. But that interest is deemed to be worth protecting because it serves the public. That is, the journalists’ interest is valued because of its usefulness to members of the public at large. The rights of priests, doctors and lawyers to preserve the confidentiality of their professional contacts are likewise justified ultimately by their value to members of the community at large” (*ibid*).

There are well-known concerns about the compatibility between this account of instrumentally valuable rights with the rest of Raz’s interest theory (Kamm 2002). Of course, if Kamm is right in her objections, the internal inconsistency of the interest theory would leave the objection against my account considered here undermotivated. Pallikkathayil (2016), at any rate, argues that Kamm’s critique fails when we attend appropriately to the distinction between intrinsically and instrumentally valuable rights. Indeed, Pallikkathayil proposes that we interpret Raz’s interest theory thus: “‘X has a right’ if and only if X can have rights, and, other things being equal, the intrinsic or instrumental value of an interest of X’s is a sufficient reason for holding some other person(s) to be under a duty” (Pallikkathayil 2016: 150). On this understanding of the interest theory, my contention in this thesis may be understood as follows: the instrumental value of the right of SARs against RSV is a sufficient reason to hold the users of SARs to be under a duty to refrain from RSV. Based on my arguments in the previous chapter, which established that RSV is a matter of concern for the basic structure of society, I believe that this contention is well-grounded.

A follow-up objection to my arguments in this section asks why, if I can convey my whole thesis in terms of derivative legal rights, I have considered it valuable and instructive to discuss legal personhood for SARs at all. Of course, the orthodox view is that all and only legal persons either hold legal rights and bear legal duties, or at least have the legal capacity to hold rights and bear

duties.¹⁰⁵ As a result, the ascription of legal personhood to SARs is necessary for the ascription to them of any legal rights. But if legal personhood is simply a shorthand for legal rights and duties and if, as I have argued, the legal rights and duties of SARs are only instrumentally valuable, did we really need all this talk of legal personhood in this chapter?¹⁰⁶ There are two reasons why we did. Firstly, the ascription of legal personhood to SARs is not just a necessary condition for the ascription of legal rights and/or duties. Rather, there is a flexibility and an open-endedness to ascriptions of legal personhood which leaves open the possibility of ascribing to the entity new legal rights and responsibilities as the function of that entity in society changes with time. This sometimes controversial possibility of broadening and deepening the nature and scope of a legal person's rights is adverted to by the recent series of US Supreme Court rulings concerning the legal rights of corporations which I alluded to in §4.1.2. I do not purport in this thesis to offer an exhaustive list of the legal rights and responsibilities of SARs and have focused only on the RSV context. As a result, the flexibility and open-endedness of the ascription of legal personhood is my invitation to future scholarship to take the project further.

Secondly, and more subtly, as I explained above, there is an important expressive function of certain legal statements which serves to shape human attitudes, beliefs and behaviours in valuable ways. In the context of the environmental movement, Stone (1972[2010]: 23-31) has adverted to the valuable “psychic and socio-psychic aspects” of ascribing legal personhood to certain kinds of natural objects, reflecting and sustaining a social norm which values the environment in its own right. Whilst ascriptions of legal personhood to new entities are often justified by appeal to the interests of existing legal persons (*ibid.* 24), the expressive function of the law means that the legal

¹⁰⁵ See *supra* fn.2.

¹⁰⁶ In his discussion of moral status (i.e. personhood), DeGrazia (2008: 184) believes that “talk of moral status is redundant and can always in principle be replaced by other language”. Stone (1972[2010]: 22) considers a similar objection to his proposal that natural objects, such as trees, should have the kind of legal status which he calls “standing”.

personhood of natural objects encourages us to interact with the natural world with care and responsibility. In a similar way, the ascription of legal personhood to autonomous robots such as SARs may encourage norms of sexual engagement which embody care and responsibility, something which may turn out to be valuable not just to women and girls but also, incidentally, as a more intrinsically valuable way of being.

§4.3. Concluding Remarks

In this chapter, I have reflected on the proposal with which I opened this thesis, namely the idea tabled by the European Parliament that at least some autonomous robots should be granted legal personhood. I have argued for a functionalist interpretation of personhood moralism (§4.1), a position which supports the ascription of legal personhood to SARs, including derivative legal rights against subjection to RSV (§4.2). My arguments in this chapter are not the end of the story but I have sought to articulate an appropriate legal response to the RSVP and, in so doing, I have offered some new reasons in support of the European Parliament's proposal to ascribe legal personhood to at least some autonomous robots.

5. Summing Up

Ending is better than mending.

(Huxley 2002: 32)

§5.1. Closing Considerations

I wish to close this thesis by considering three questions which my arguments so far might evoke. The answers which I offer to these questions make no pretence to offer closure. But the thesis would seem wanting were I to offer no reflections on these matters.

Firstly, there is another kind of sexual violence which the emergence of SARs will make possible, one which reverses the roles played by human and SAR in the RSV cases which I have focused on hitherto. For example, the development of “male” SARs would make possible the following, disturbing case:

Reverse Robotic Rape: A tried-and-tested SAR, B, penetrates with its “penis” the vagina anus or mouth of a human, A. A has not indicated consent. This is not a mechanical malfunction and no human intended or foresaw it; rather, it results from the SAR’s machine learning.¹⁰⁷

It is clear that, had a human committed the act, they would be charged with rape. The distinctive problem with machine learning systems such as SARs is that there appears to be a morally problematic “responsibility gap” when these systems cause wrongful harms for which no person seems to bear legal responsibility.¹⁰⁸ Some have argued that all those who contribute to the

¹⁰⁷ To illustrate: one deep-learning system was trained to reliably distinguish pictures of wolves from pictures of huskies; when it misclassified very clear images, it became apparent that the system had learnt to use the presence of snow in the picture as a determining factor (Ribeiro *et al.* 2016). Likewise, SARs might be overwhelmingly successful at recognizing signs of consent but for the wrong reasons, such that they go wrong in surprising cases.

¹⁰⁸ This kind of responsibility gap is what Santoni de Sio and Mecacci (2021) call an “attributability gap”. Cf. *supra* fn. 4.

production of the autonomous system should bear strict criminal liability for even unforeseeable damages which the system might cause (Burri 2017; Heyns 2013; Corn 2014). Another solution to responsibility gaps in the AI ethics literature is to introduce faultless civil law compensation schemes for damages caused by AI systems. Schellekens (2018) believes that this could replace traditional liability schemes for self-driving vehicles (cf., Crootof 2016).¹⁰⁹ There are concerns, however, that either of these schemes would have a “chilling effect” on AI innovation (Trudy *et al.* 2021: 272). A third solution, as I explained at the start of this thesis, is to introduce legal personhood for AI systems, à la *European Parliament resolution on Civil Law Rules on Robotics*, “so that at least the most sophisticated autonomous robots could [be legally] responsible for making good any damage they may cause” (Delvaux 2017: 18). My arguments in the previous chapter might seem to support this latter approach to the regulation of cases such as Reverse Robotic Rape.¹¹⁰

Secondly, then, you might be wondering how this kind of legal personhood for SARs would play out in practice. In his (2019), Turner glosses a variety of possibilities but the most promising one appears to be to house the legal personhood of AI systems like SARs within existing legal structures governing corporations. Bayern (2016), for example, has argued that the US law of limited liability companies (LLCs) could be utilised to bestow legal personhood on any autonomous systems. In principle, he shows, it is possible to create an LLC whose operating agreement placed it under the control of an AI system and then have every other member of the LLC withdraw, leaving the system unsupervised by humans. In practice, there may be problems with housing AI legal personhood within existing legal frameworks given that this would often not be in harmony with legislative intention (Scherer 2017). But the attractions of the principle are manifold: as Turner

¹⁰⁹ Santoni de Sio and Mecacci (2021: 1078) criticise this approach. The context of self-driving vehicles generates legal controversy. Recently, when a self-driving car ran over a pedestrian, the onboard operator was considered negligent *and* liable for the death (BBC 2020). This approach strikes many as mistaken (Bellon 2018; Santoni de Sio and Mecacci 2021: 1078; Stilgoe 2020).

¹¹⁰ For one account of how autonomous robots could, at least in principle, be *punished* in criminal law: Abbott and Sarch (2019).

(2019: 187-188) explains, treating AI systems rather like LLCs would not only encourage innovation—through acting as a valuable “firewall” between existing legal persons and damages potentially caused by autonomous systems—but also allow AI systems to hold property, something which would resolve issues raised concerning the ownership of new intellectual property of which AI is the creator.¹¹¹

Finally, some might be curious as to how the development of general AI might alter the kind of legal framework which I have sketched in this thesis. As I explained in chapter three, general AI would display humanlike intelligence and, as a result, would not be that motor for the replication and entrenchment of existing bias which is involved in narrow AI-based autonomy. But equally, as I explained in chapter four, general AI would seem to possess at least some of the features commonly associated with *intrinsic* moral status, including sentience. As a result, general AI would appear to be the appropriate subject of non-derivative moral and legal rights. Ascriptions of legal personhood to SARs, then, seems to lay regulatory foundations appropriate for the open-ended development of more sophisticated AI systems, to which would accrue more legal rights as their relationships with humans become more complex. At the same time, however, given their seeming intrinsic moral status, the production, sale and use of SARs endowed with general AI might appear to be radically immoral, more like sex trafficking than pornography. In other words, whilst the ascription of legal personhood would be a core part of developing an appropriate legal framework for general AI SARs as well as narrow AI ones, the roles which general AI could appropriately occupy in the sphere of social robotics might turn out to be more limited than the roles appropriate to narrow AI.

¹¹¹ Bryson *et al.* (2017: 4.2.1) contend that ascriptions of legal personhood to AI systems would lead to “natural persons using an artificial person to shield themselves from the consequences of their conduct”. Turner (2019: 191) points out that this line of argument is not just speculative but moreover entails the absurd conclusion that we should abolish “limited liability”—and, thus, companies—altogether.

§5.2. Conclusion

In conclusion, I have sought to make two principal contributions to the AI ethics literature in this thesis. Firstly, I have articulated and defended a new argument for the legal proscription of RSV, one which appeals to politically liberal principles which all free and equal citizens in a pluralistic society can reasonably accept. This line of argument is dialectically important because, until now, arguments for the legal proscription of RSV have depended upon the invocation of moral conceptions of the good, whilst politically liberal principles have seemed to support the legal permissibility of RSV.

Secondly, I have sought to shift the focus in the academic and public debate surrounding the ascription of legal personhood to autonomous robots. Legal personhood has often been touted as a device which will assist us in developing an appropriate legal framework to regulate wrongs caused *by* autonomous robots. But I have argued in this thesis that legal personhood may also assist us in developing an appropriate legal framework to regulate wrongs inflicted *on* autonomous robots. This is vitally important: what we owe to autonomous robots, both morally and legally, is a matter which we as private individuals and public legislators can no longer afford to ignore.

In sum, given the imminent emergence of autonomous robots in spheres both public and private, a legal framework for their regulation is urgently needed. This thesis is my contribution to that wider project.

Bibliography

Abbey, R., 2011, *The Return of Feminist Liberalism*, Taylor & Francis Group

Abbott, R., and Sarch, A., 2019, “Punishing Artificial Intelligence: Legal Fiction and Science Fiction”, *University of California, Davis* 53: 323

Ackerman, B., 1980, *Social Justice and the Liberal State*, New Haven: Yale University Press

Adarsh, H. and Sahoo, S., 2023, “Pornography and Its Impact on Adolescent/Teenage Sexuality”, *Journal of Psychosexual Health* 5(1)

Adler, P., Falk, C., Friedler, S. A., Rybeck, G., Scheidegger, C., Smith, B., Venkatasubramanian, S., 2016, “Auditing Black-box Models for Indirect Influence”, IEEE International Conference on Data Mining (ICDM), 2016, <https://arxiv.org/abs/1602.07043>

Altwicker, T., 2015, “Rechtsperson im Rechtspositivismus” in R. Gröschner, S. Kirste, O. Lembcke (eds.), *Person und Rechtsperson: zur Ideengeschichte der Personalität*, Mohr Siebeck

Amoroso, D. and Tamburrini, G., 2021, “Toward a Normative Model of Meaningful Human Control over Weapons Systems, *Ethics & International Affairs* 35, no. 2: 245-272

Anderson, E., 1991, “John Stuart Mill and Experiments in Living”, *Ethics* 102: 4–26

Anderson, M., and Anderson, S., 2011, “General Introduction,” in M. Anderson and S. Anderson (eds.) *Machine Ethics*, Cambridge: Cambridge University Press: 1–4

Anthropic, 2023, “Claude’s Constitution”, May 9 2023, <https://www.anthropic.com/index/claudes-constitution>

Arkin, R., 2009, *Governing Lethal Behavior in Autonomous Robots*, Boca Raton: CRC Press

- Arkoudas, K., Bringsjord, S. and Bello, P., 2005, “Toward ethical robots via mechanized deontic logic”, in *Machine ethics: Papers from the 2005 AAAI fall symposium*: 17-23
- Asaro, P., 2012, On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making, 94 INT’L REV. RED CROSS 687, 701 (2012);
- , 2016, “Jus nascendi, Robotic Weapons and the Martens Clause” in R. Calo, M. Froomkin and I. Kerr (eds.) *Robot Law*, Cheltenham UK: Edward Elgar Publishing: 367-386
- , 2020, “Autonomous Weapons and the Ethics of Artificial Intelligence”, in S. M. Liao (ed.) *Ethics of Artificial Intelligence*, Oxford University Press 2020
- Asimov, I., 1950, “Runaround”, reprinted in *I, Robot*, New York: Doubleday
- Audi, R., 2006, “Testimony, credulity, and veracity”, *The Epistemology of Testimony*, J. Lackey and E. Sosa (eds.), Oxford: Oxford University Press: 25–49
- Austrian Civil Code, Available online: https://www.trans-lex.org/602100/_/austrian-civil-code/
- Baehr, A., 2008, “Perfectionism, Feminism and Public Reason”, *Law and Philosophy* 27: 193-222
- Bartneck, C., van der Hoek, M., Mubin, O., and Al Mahmud, A., 2007, “Daisy, Daisy, Give me your answer do!” – *Switching off a robot*, PROCEEDINGS OF THE 2ND ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION, Washington DC: 217–22
- Bayern, S., 2016, “The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems”, *European Journal of Risk Regulation* 2: 297-309
- BBC, 2020, “Uber’s self-driving operator charged over fatal crash” <https://www.bbc.com/news/technology-54175359>. Accessed 8 April 2021

BBC, 2021, “Tesla Autopilot: US opens official investigation into self-driving tech”,
<https://www.bbc.co.uk/news/technology-58232137>

Bellon, T., 2018, “Fatal U.S. self-driving auto accident raises novel legal questions” *Reuters*. <https://www.reuters.com/article/us-autos-selfdriving-uber-liability-anal/fatal-u-s-self-driving-auto-accident-raises-novel-legal-questions-idUSKBN1GW2SP>. Accessed 8 April 2021

Benhabib, S., 2002, *The Claims of Culture: Equality and Diversity in the Global Era*, Princeton: Princeton University Press

Berg, J., 2007, “Of Elephants and Embryos: A Proposed Framework for Legal Personhood”, 59 *Hastings Law Journal*

Berk, R. A. and Bleich, J., 2013, “Statistical Procedures for Forecasting Criminal Behaviour: A Comparative Assessment”, *Criminology & Public Policy* 12(3)

Bianchi, A., 2020, “Considering sex robots for older adults with cognitive impairments”, *Journal of Medical Ethics* Vol. 47, No. 1

Bilchitz, D., 2009, “Moving Beyond Arbitrariness: The Legal Personhood and Dignity of Non-Human Animals”, *South African Journal on Human Rights* 25(1): 38-72

Billingham, P. and Taylor, A., 2023, “Can Civic Friendship Ground Public Reason?”, *The Philosophical Quarterly* Vol 0, No. 0: 1-22

Bird, C., 2014, “Coercion and Public Justification,” *Politics, Philosophy, & Economics*, 13(3): 189–214

Blackstone, W., 1876, “The commentaries on the laws of England of Sir William Blackstone”,
London : John Murray, Available online:
<https://archive.org/details/commentariesonla01blacuoft/page/viii/mode/2up>

Bohman, J., 1996, *Public Deliberation: Pluralism, Complexity, and Democracy*, Cambridge, MA: MIT Press

—, 2003, “Deliberative Toleration,” *Political Theory*, 31(3): 757–779

Bostrom, N., 2012, “The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents”, *Minds and Machines* 22 no. 2: 71-85

—, 2014, *Superintelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press

Braithwaite, S. R., Coulson, G., Keddington, K. and Fincham, F. D., 2015, “The Influence of Pornography on Sexual Scripts and Hooking Up Among Emerging Adults in College”, *Archives of Sexual Behavior*, 44: 111–23

Brennan, S., 2004, “Liberalism, Feminism, and Rights”, in *Varieties of Feminist Liberalism*, A. Baehr (ed.), Lanham MD: Rowman and Littlefield

Brennan, T. and Oliver, W. L., 2013, “The Emergence of Machine Learning Techniques in Criminology”, *Criminology & Public Policy*, 12(3)

Brennan, J. and Jaworski, M., 2015, “Markets Without Symbolic Limits”, *Ethics* 125(4): 1053-1077

Brink, D., 2008, “Mill’s Liberal Principles and Freedom of Expression” in *Mill’s ‘On Liberty’: A Critical Guide*, C. L. Ten (ed.), New York: Cambridge University Press

Brown, C., 2019, “Sex Robots, Representation, and the Female Experience”, *The American Papers* 37: 105–118

<https://amst.fullerton.edu/students/Final%20AP%202019%20PMG%20PRESS.pdf#p%20age=10>

[5](#)

- Bryson, J. J., Diamantis, M. E., & Grant, T. D., 2017, “Of, for, and by the people: the legal lacuna of synthetic persons”, *Artificial Intelligence and Law*, 25(3), 273–291. <https://doi.org/10.1007/s10506-017-9214-9>
- Buchwald, E., Fletcher, P. R., Roth, M. (eds.), 2005, *Transforming a rape culture*, Milkweed editions, Minneapolis
- Buhbut, A., 2021, “בלעדִי | להקות הרחפנים של צה"ל השתתפו בלחימה בעזה, וחוקי המשחק צפויים להשתנות”, Walla News June 5 2021, <https://news.walla.co.il/item/3439695>
- Burge, T., 1993, “Content preservation”, *The Philosophical Review* 102: 457–488
- , 1997, “Interlocution, perception, and memory”, *Philosophical Studies* 86: 21–47
- Burge-Hendrix, B., 2007, “The Educative Function of Law”, *Law and Philosophy* M. Freeman and R. Harrison (eds.), OUP: 243-254
- Burri, S., 2017, “What is the Moral Problem with Killer Robots?”, in B. J. Strawser, R. Jenkins, and M. Robillard (eds.) *Who Should Die? The Ethics of Killing in War*, Oxford Scholarship Online
- Burnwell v. Hobby Lobby Stores, Inc.* [2014] 573 U.S. 682
- Campbell, J., 2020, *Causation in Psychology*, Harvard University Press
- Carline, A., 2011, “Criminal justice, extreme pornography and prostitution: Protecting women or promoting morality?”, *Sexualities* 14(3): 312-333
- Carpenter, J., 2013, *THE QUIET PROFESSIONAL: AN INVESTIGATION OF U.S. MILITARY EXPLOSIVE ORDNANCE DISPOSAL PERSONNEL INTERACTIONS WITH EVERYDAY FIELD ROBOTS*, Dissertation, University of Washington

CEDAW, 1979, “Convention on the Elimination of All Forms of Discrimination against Women”, Available online: <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-elimination-all-forms-discrimination-against-women>

Chambers, C., 2018, *Against Marriage: An Egalitarian Defense of the Marriage-Free State*, Oxford: Oxford University Press

Chatterjee, B. B., 2020, “Child sex dolls and robots: challenging the boundaries of the child protection framework”, *International Review of Law, Computers & Technology*, Vol. 34, No.1: 22-43

Chesterman, S., 2020, “Artificial Intelligence and the Limits of Legal Personality”, *International & Comparative Law Quarterly* 69(4): 819-844

Chopra, S. & White, L. F., 2011, *A legal theory for autonomous artificial agents*, Ann Arbor: University of Michigan Press

Citizens United v. Federal Election Commission [2010] 558 U.S. 310

Coffee, J. C., Jr., 1991, Does “Unlawful” Mean “Criminal”? Reflections on the Disappearing Tort/Crime Distinction in American Law, 71 B.U. L. REV. 193

Cohen, J., ms, “Don’t Shoot the Algorithm”, personal communication

Commonwealth of Australia, 2009, *Australian Government Attorney-General Department Discussion Paper: “Should the Australian National Classification Scheme include an R18+ classification category for computer games?”*, Attorney General’s Department, Canberra

Conard, A., 1976, *Corporations in Perspective*, Foundation Press

Corn, G., 2014, Autonomous Weapon Systems: Managing the Inevitability of “Taking the Man Out of the Loop” 21 (unpublished manuscript), <https://www.law.upenn.edu/live/files/3894-corn-understanding-the-loop-regulating-the-next> [<https://perma.cc/9KMQ-2VGQ>].

Cowen, N., 2015, “Millian Liberalism and Extreme Pornography”, *American Journal of Political Science* 60(2): 509-520

CPS (Crown Prosecution Service), 2021, “Tackling Rape Myths and Stereotypes”, in *Legal Guidance, Rape and Sexual Offences*: chapter 4. Available online: <https://www.cps.gov.uk/legal-guidance/rape-and-sexual-offences-chapter-4-tackling-rape-myths-and-stereotypes>

Crépault, C. and Couture, M., 1980, “Men’s Erotic Fantasies”, *Archives of Sexual Behaviour* 9

Criminal Justice and Immigration Act, 2008, available online: <https://www.legislation.gov.uk/ukpga/2008/4/part/5/crossheading/pornography-etc>

Critelli, J. and Bivona, J., 2008, “Women’s Erotic Rape Fantasies: An Evaluation of Theory and Research”, *Journal of Sex Research* 45

Crootof, R., 2015, “The Killer Robots Are Here: Legal and Policy Implications”, 36 *CARDOZO L. REV.* 1837

—, 2016, “War Torts: Accountability for Autonomous Weapons”, *University of Pennsylvania Law Review*, May 2016, Vol. 164, No. 6: 1347-1402

Danaher, J., 2017a, “Robotic Rape and Robotic Child Sexual Abuse: Should They be Criminalised?”, *Criminal Law and Philosophy* 11: 71-95

—, 2017b, “The Symbolic Consequences Argument in the Sex Robots Debate”, in J. Danaher and N. McArthur (eds.) *Robot Sex: Social and Ethical Implications*, MIT Press, Cambridge

——, 2019, “Regulating Child Sex Robots: Restriction or Experimentation?”, *Med Law Rev.* 2019 Nov 1, 27(4): 553-575

Darling, K., 2014[2016], “Extending legal protection to social robots: the effects of anthropomorphism, empathy, and violent behavior towards robotic objects” in M. Froomkin, R. Calo, I. Kerr (eds.), *Robot Law*, Edward Elgar, Cheltenham

Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2044797

Dartmouth College v Woodward [1819] 17 US 518

de Beauvoir, S., 1956, *The Second Sex*, H. M. Parshley (trans., ed.), Jonathan Cape: Thirty Bedford Square London

de las Heras Ballell, T., 2010, *Introduction to Spanish Private Law: Facing the Social and Economic Challenges* Routledge-Cavendish

DeGrazia, D., 2008, “Moral Status As a Matter of Degree?”, *The Southern Journal of Philosophy* Vol. XLVI: 181-198

Deep Learning Glossary, Deep AI: <https://deepai.org/machine-learning-glossary-and-terms/deep-learning>

Delvaux, M. 2017, Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), *Committee on Legal Affairs*, <https://www.europarl.europa.eu/cmsdata/113782/juri-final-report-robotics.pdf>

Dennett, D., 1987, *The Intentional Stance*, MIT Press

——, 2009, “Intentional systems theory”, in A. Beckermann, B. McLaughlin, & S. Walter (eds.), *The Oxford Handbook of Philosophy of Mind*, Oxford University Press: 339-350

Devlin, K., 2020, “Meet Samantha, The Sex Robot That Gives You Life Advice”, *Refinery 29*

<https://www.refinery29.com/en-gb/sex-robots-kate-devlin>

——, 2018, *Turned on: Science, sex and robots*, London: Bloomsbury Publishing

Devlin, P., [1965]1968, *The Enforcement of Morals*, Oxford: Oxford University Press

Dewey, J., 1926, “The Historic Background of Corporate Legal Personality”, 35 *Yale Law Journal*

Donner, W., 2008, “Autonomy, Tradition, and the Enforcement of Morality”, in C. L. Ten (ed.) *Mill’s ‘On Liberty’: A Critical Guide*, New York: Cambridge University Press

Dudgeon v. The United Kingdom [1981] ECHR 22 Oct 1981

Duff, A., 2014, “Towards a Modest Legal Moralism”, *Criminal Law and Philosophy* 8: 217-235

——, 2018, *The Realm of Criminal Law*, OUP

Duffy, B., 2003, *Anthropomorphism and the social robot*, ROBOTICS AND AUTONOMOUS SYSTEMS 42: 179-83

Dworkin, R., 1978, “Liberalism,” in *Public and Private Morality*, S. Hampshire (ed.), Cambridge: Cambridge University Press, pp. 113–43.

——, 1986, *Law’s Empire*, Belknap Press

——, 1995, “Foundations of Liberal Equality,” in *Equal Freedom*, S. Darwall (ed.), Ann Arbor: University of Michigan Press.

——, 2011, *Justice for Hedgehogs*, Cambridge, Mass.: Harvard University Press

Dyzenhaus, D., 1992, “John Stuart Mill and the Harm of Pornography”, *Ethics* 102: 534-551

Dwyer, S. (ed.), 1995, *The Problem of Pornography*, Belmont CA: Wadsworth

E&T editorial staff, 2017, “Social implications of sex robots scrutinised in new report”, July 5 2017, <https://eandt.theiet.org/content/articles/2017/07/social-implications-of-sex-robots-scrutinised-in-new-report/>

Ebels-Duggan, K., 2010, “The Beginning of Community: Politics in the Face of Disagreement,” *The Philosophical Quarterly*, 60(283): 50–71

Ecuador Constitution, Available online: https://www.constituteproject.org/constitution/Ecuador_2021?lang=en

Edmundson, W. A., 2004, *An Introduction to Rights*, Cambridge Introductions to Philosophy and Law), second edition, Cambridge/New York: Cambridge University Press; second edition, 2012

Eidenmuller, H. and Wagner, G., 2021, *Law by Algorithm*, Mohr Siebeck

Einsiedel, E., 1992, “The Experimental Research Evidence; Effects of Pornography on the ‘Average Individual’”, in C. Itzin (ed.) *Pornography: Women , Violence and Civil Liberties*, Oxford: Oxford University Press: 248-283

Eliot, T. S., 1921, review of *Metaphysical Lyrics and Poems of the Seventeenth Century: Donne to Butler*. Selected and edited, with an Essay, by Herbert J. C. Grierson (Oxford: Clarendon Press. London; Milford) in the *Times Literary Supplement*, October 1921

Eskens, R., 2017, “Is Sex With Robots Rape?”, *Journal of Practical Ethics*: 62-76

European Economic and Social Committee, 2017, “Opinion on ‘Artificial Intelligence—The Consequences of Artificial Intelligence on the (Digital) Single Market, Production, Consumption, Employment and Society’ (Own-Initiative Opinion) OJ C288/1”

European Parliament Resolution 2020/2014[2021], European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)),

<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020IP0276>

Feinberg, J., 1984, *Harm to others*, New York: Oxford University Press

Feldman, M., Friedler, S., Moeller, J., Scheidegger, C., Venkatasubramanian, S., 2015, “Certifying and removing disparate impact”, 2015 ACM SIGKDD Conference on Knowledge Discovery and Data Mining

<https://arxiv.org/abs/1412.3756>

Ferguson, C. J., & Hartley, R. D., 2022, “Pornography and Sexual Aggression: Can Meta-Analysis Find a Link?”, *Trauma, Violence, & Abuse*, 23(1): 278–287

Finnis, J., 1981, *Natural law and natural rights*, 1st ed., Oxford: Oxford University Press

——, 1997a, “The Good of Marriage and the Morality of Sexual Relations: Some Philosophical and Historical Observations”, *American Journal of Jurisprudence* 42

——, 1997b, “Law, Morality, and ‘Sexual Orientation’”, in J. Corvino (ed.), *Same Sex: Debating the Ethics, Science, and Culture of Homosexuality*, Rowman and Littlefield

——, 2008, “Marriage: A Basic and Exigent Good”, *The Monist*, Vol. 91, Nos. 3 & 4: 388-406

——, 2011a, *Natural law and natural rights*, 2nd ed. Oxford: Oxford University Press

——, 2011b, “The Priority of Persons”, in J. Finnis, *Intention and Identity*, OUP

——, 2013, “The Priority of Persons Revisited”, *The American Journal of Jurisprudence* 58(1): 45-62

- Fitzgerald, P. J., 1966, *Salmond on Jurisprudence*, twelfth edition, London: Sweet & Maxwell
- Forrest, K. B., 2021, “When Machines Can Be Judge, Jury, and Executioner: Justice in the Age of Artificial Intelligence”, *World Scientific Publishing Co. Pte. Ltd.*
- Frank, L. and Nyholm, S., 2017, “Robot sex and consent: Is consent to sex between a robot and a human conceivable, possible, and desirable?”, *Artificial Intelligence Law* 25: 305-323
- Freund, E., [1897]2000, *The Legal Nature of Corporations*, Batoche Books
- Fricker, M., 2003, “Epistemic Injustice and a Role for Virtue in the Politics of Knowing”, *Metaphilosophy*, 34(1–2): 154–173
- , 2007, *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford: Oxford University Press
- Friedman, B., Kahn, P., Hagman, J., 2003, “Hardware Companions? – What Online AIBO Discussion Forums Reveal about the Human-Robotic Relationship”, *PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS*: 273–80
- Fritz, N. and Paul, B., 2017, “From Orgasms to Spanking: A Content Analysis of the Agentic and Objectifying Sexual Scripts in Feminist, for Women, and Mainstream Pornography”, *Sex Roles* 77: 639–52
- Frydrych, D., 2018, “State-of-the-Art: The Theories of Rights Debate” (February 13, 2018), *Jurisprudence: An International Journal of Legal and Political Thought* 9(3): 566-588
- Gagnon, J. H. and Simon, W., 1973, *Sexual Conduct: The Social Sources of Human Sexuality*, Aldine
- Gaita, R., 1999, *A Common Humanity: Thinking About Love & Truth & Justice*, Text Publishing, Melbourne

- Galaitsi, S. E., Hendren, C. O., Trump, B. and Linkov, I., 2019, “Sex Robots—A Harbinger for Emerging AI Risk”, *Frontiers of Artificial Intelligence*
- Galliot, J., 2015, “Military Robots: Mapping the Moral Landscape” Surrey UK: Ashgate
- Galliot, J., D. MacIntosh and J.D. Ohlin (eds.), 2021, *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*, Oxford University Press
- Gardner, J., and Shute, S., 2000, “The Wrongness of Rape”, in J. Horder (eds.) *Oxford Essays in Jurisprudence*, Oxford University Press
- Garreau, J., 2007, “Bots on The Ground in the Field of Battle (Or Even Above It), Robots Are a Soldier’s Best Friend”, *WASHINGTON POST* (May 6, 2007)
- Gaus, G., 1996, *Justificatory Liberalism: An Essay on Epistemology and Political Theory*, Oxford: Oxford University Press
- , 2009, “The Place of Religious Belief in Public Reason Liberalism,” in *Multiculturalism and Moral Conflict*, M. Dimova-Cookson and P. Stirk (eds.), New York: Routledge, pp. 19–37
- , 2011, *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*, Cambridge: Cambridge University Press
- , 2012, “Sectarianism Without Perfection? Quong’s Political Liberalism,” *Philosophy and Public Issues*, 2(1): 7–15
- , 2016, *The Tyranny of the Ideal*, Princeton, NJ: Princeton University Press
- , 2017, “Self-Organizing Moral Systems,” *Politics, Philosophy, & Economics*, first published online 29 August 2017. doi: 10.1177/1470594X17719425

- Gaus, G. and K. Vallier, 2009, "The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Asymmetry, and Political Institutions," *Philosophy & Social Criticism*, 35(1): 51–76
- George, R., 2004, "What's Sex Got To Do With It? Marriage, Morality, and Rationality", *The American Journal of Jurisprudence* 49(1): 63-85
- German Civil Code, *Bürgerliches Gesetzbuch*, BGBI, Available online: <https://www.gesetze-im-internet.de/bgb/BGB.pdf>
- Gersen, J. S., 2019, "Sex Lex Machina", *Columbia Law Review* Vol. 199, No. 7: 1793-1810
- Gips, J., 1995, "Towards the Ethical Robot", Second International Workshop on Human & Machine Cognition, Perdido Key, Florida, 1991, published in *Android Epistemology*, K. Ford, C. Glymour, P. Hayes (eds.), MIT Press
- Goldberg, S., 2010, *Relying on Others: An Essay in Epistemology*, New York: Oxford University Press
- , 2013, "Anonymous assertions", *Episteme* 10(2): 134–151
- , 2015, *Assertion: On the Philosophical Significance of Assertoric Speech*, Oxford: Oxford University Press
- Golding, M. P. and W. A. Edmundson (eds.), 2005, *The Blackwell Guide to the Philosophy of Law and Legal Theory*, Oxford: Blackwell Publishing
- Gray, J. C., 1996, *Mill on Liberty: A Defence*, 2nd ed., New York: Routledge
- , 1997, *The Nature and Sources of the Law*, D. Campbell and P. Thomas (eds.), Ashgate
- Green, L., 2013, "Should Law Improve Morality?", *Criminal Law and Philosophy* 7: 473-494

Greshko, M., 2018, “Meet Sophia, the Robot That Looks Almost Human”, *National Geographic*

<https://www.nationalgeographic.co.uk/science-and-technology/2018/05/meet-sophia-the-robot-that-looks-almost-human>

Griswold v. Connecticut [1965] 381 U.S. 479

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

Gutiu, S., 2012, “Sex Robots and the Roboticization of Consent”, *We Robot Law Conference Miami*

——, 2016, “The Roboticization of Consent”, in *Robot Law*, Calo, Froomkin and Kerr (eds.), Cheltenham, UK; Edward Elgar Publishing

Habermas, J., 1990, *Moral Consciousness and Communicative Action*, C. Lenhardt and S. W. Nicholsen (trans.), Cambridge, MA: MIT Press

——, 1995, “Reconciliation Through the Public use of Reason: Remarks on John Rawls’s Political Liberalism,” *The Journal of Philosophy*, 92(3): 109–131

——, 1996, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, W. Rehg (trans.), Cambridge, MA: MIT Press

——, 1998, *The Inclusion of the Other: Studies in Political Theory*, C. Cronin and HP. DeGreiff (eds.), Cambridge, MA: MIT Press

Hallis, F., 1930, *Corporate Personality. A Study in Jurisprudence*, OUP

Hampton, J., 1984, “The Moral Education Theory of Punishment”, *Philosophy & Public Affairs*, Vol. 13, No. 3: 208-238

Hänel, H., 2021, “Who’s to Blame? Hermeneutical Misfire, Forward-Looking Responsibility, and Collective Accountability”, *Social Epistemology* 35(2): 173-184

Hanson Robotics, 2023, <https://www.hansonrobotics.com/sophia/>

Harel, A., 2005, “Theories of Rights”, in Golding and Edmundson 2005: 191–206

Hart, H., [1964]1968, *Law, Liberty and Morality*, Oxford: Oxford University Press

——, 1982, *Essays on Bentham: Studies in Jurisprudence and Political Theory*, Oxford: Clarendon Press

——, 1983, *Essays in Jurisprudence and Philosophy*, Oxford: Oxford University Press

Hartley, C. and L. Watson, 2010, “Is a Feminist Political Liberalism Possible?” *Journal Of Ethics & Social Philosophy*, 5(1): 1–21

Haslanger, S., 2005, “What Are We Talking About? The Semantics and Politics of Social Kinds”, *Hypatia*, 20(4): 10–26

Hegel, F., Muhl, C., Wrede, B., Hielscher-Fastabend, M., and Sagerer, G., 2009, *Understanding Social Robots*, The Second International Conferences on Advances in Computer-Human Interactions: 169-74.

Herman, D., 1988, “The rape culture”, *Culture* 1(10): 45-53

Hern, A., 2017, “Give robots ‘personhood’ status, EU committee argues”, *The Guardian*

Herzfeld, N., 2017, “Religious Perspectives on Sex with Robots”, in J. Danaher and N. McArthur (eds.) *Robot Sex: Social and Ethical Implications*: 122-136

Heyns, C., 2013, UN Human Rights Council, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, Lethal Autonomous Robotics, A/HRC/23/47,

April

9,

2013,

http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf

Hill, K., 2014, “Are Child Sex-Robots Inevitable?”, *Forbes*

<https://www.forbes.com/sites/kashmirhill/2014/07/14/are-child-sex-robots-inevitable/?sh=529beb0e460b>

Hines, D., 2009, founder of TrueCompanion, cited in [https://techcrunch.com/2010/01/09/nsfw-true-companion-debuts-sex-robot-roxxy/?guccounter=1&guce_referrer=aHR0cHM6Ly9kdWNrZHVja2dvLmNvbS8&guce_referrer_sig=AQAAANVNI210M_gyuakfa2Ua14z6F-Jq5Y6-](https://techcrunch.com/2010/01/09/nsfw-true-companion-debuts-sex-robot-roxxy/?guccounter=1&guce_referrer=aHR0cHM6Ly9kdWNrZHVja2dvLmNvbS8&guce_referrer_sig=AQAAANVNI210M_gyuakfa2Ua14z6F-Jq5Y6-GGNgVtit5sAb1szozbrprKWuQcji3dypMnITZBriRuXZmvX7QRUAbp882P7I8jWLBlikACnYrBlKelNTV_aM31c27ubi1WmS_j67VaQCCPXBOuAtmnhA5ORCfQgD4UVbonMphcd2vl5l)

[GGNgVtit5sAb1szozbrprKWuQcji3dypMnITZBriRuXZmvX7QRUAbp882P7I8jWLBlikACnYrBlKelNTV_aM31c27ubi1WmS_j67VaQCCPXBOuAtmnhA5ORCfQgD4UVbonMphcd2vl5l](https://techcrunch.com/2010/01/09/nsfw-true-companion-debuts-sex-robot-roxxy/?guccounter=1&guce_referrer=aHR0cHM6Ly9kdWNrZHVja2dvLmNvbS8&guce_referrer_sig=AQAAANVNI210M_gyuakfa2Ua14z6F-Jq5Y6-GGNgVtit5sAb1szozbrprKWuQcji3dypMnITZBriRuXZmvX7QRUAbp882P7I8jWLBlikACnYrBlKelNTV_aM31c27ubi1WmS_j67VaQCCPXBOuAtmnhA5ORCfQgD4UVbonMphcd2vl5l)

Horowitz, M. and Scharre, P., 2015, “An Introduction to Autonomy in Weapon Systems”, unpublished working paper,

https://www.files.ethz.ch/isn/188865/Ethical%20Autonomy%20Working%20Paper_021015_v02.pdf

Hohfeld, W. N., 1909, “Nature of Stockholders’ Individual Liability for Corporation Debts”, *Columbia Law Review* 9(4): 285-320

——, 1919, *Fundamental Legal Conceptions*, Walter Wheeler Cook (ed.), New Haven, CT: Yale University Press

Horwitz, M., 1985, “Santa Clara Revisited: The Development of Corporate Theory”, *West Virginia Law Review* 88

Husak, D., 2007, *Overcriminalization: The limits of the criminal law*, Oxford: Oxford University Press

Huxley, A., 2002, *Brave New World*, Available online: <https://plato-philosophy.org/wp-content/uploads/2016/05/BraveNewWorld-1.pdf>

Ishiguro, K., 2021, *Klara and the Sun*, Faber & Faber

Italian Civil Code, Available online: https://www.trans-lex.org/601300/_/italian-codice-civile/

Jecker, N., 2020, “Nothing to be ashamed of: sex robots for older adults with disabilities”, *Journal of Medical Ethics* 47: 26-32

Jenkins, H., 2012, “The war between effects and meaning: rethinking the video game violence debate”, in D. Buckingham, R. Willett (eds.) *Digital generations: children, young people, and new media*, Lawrence Erlbaum Associates, Mahwah: 19-32

Jones, P., 1994, *Rights*, New York: St. Martin’s Press

——, 1999, “Human Rights, Group Rights, and Peoples' Rights”, *Human Rights Quarterly* 21(1): 80-107

Kahn, P., Kanda, T., Ishiguro, H., Freier, N., Severson, R., Gill, B., Ruckert, J., Shen, S., 2012, “‘Robovie, you’ll have to go into the closet now’: Children’s social and moral relationships with a humanoid robot”, *DEVELOPMENTAL PSYCHOLOGY*, 48(2): 303-14

Kamm, F., 2002, KAMM, *Rights*, in OXFORD HANDBOOK OF JURISPRUDENCE AND PHILOSOPHY OF LAW 484, S. J. Schapiro & J. Coleman, (eds.)

Karner, E., 2019, “Liability for Robotics: Current Rules, Challenges, and the Need for Innovative Concepts” in S. Lohsse, R. Schulze and D. Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things*, Nomos 2019

- Kelsen, H., 1919, “Zur Theorie der juristischen Fiktionen. Mit besonderer Berücksichtigung von Vaihingers Philosophie des Als Ob”, *Annalen der Philosophie* 1: 630-658
- , 1923, *Hauptprobleme der Staatsrechtslehre: entwickelt aus der Lehre vom Rechtssatze* (2nd edn.), JCB Mohr
- , 1925, *Allgemeine Staatslehre*, Julius Springer
- , 1945, *General Theory of Law and State*, Harvard University Press
- , 1991, *General Theory of Norms*, M. Harntney (trans.), Clarendon Press
- , 1992, *Introduction to the Problems of Legal Theory*, B.L. Paulson and S.L. Paulson (trans.), Oxford: Clarendon Press
- , 2005, *Pure Theory of Law*, M. Knight (tr., ed.), The Lawbook Exchange, Ltd.
- , 2006, *General Theory of Law and State*, Transaction Publishers
- , 2015, “On the Theory of Juridic Fictions. With Special Consideration of Vaihinger’s Philosophy of the As-If”, in M. Del Mar and W. Twining (eds.), C. Kletzer (trans.), *Legal Fictions in Theory and Practice* 110, Springer International Publishing
- Kiener, M., 2022, “Are We Heading Towards a Post-Responsibility Era? Artificial Intelligence and the Future of Morality”, *Practical Ethics in the News*
- Kirkpatrick, K., 2017-01-23, “It's not the algorithm, it's the data”, *Communications of the ACM* 60(2): 21–23
- Kirste, S., 2000, “Verlust und Wiederaneignung der Mitte—zur juristischen Konstruktion der Rechtsperson”, *Evangelische Theologie* 60(25)

Kleeman, J., 2017, “The race to build the world’s first sex robot”, *The Guardian*,
<https://www.theguardian.com/technology/2017/apr/27/race-to-build-world-first-sex-robot>

——, 2020, *Sex Robots & Vegan Meat*, Picador

Klosko, G., 2003, “Reasonable rejection and neutrality in justification”, in *Perfectionism and neutrality: Essays in liberal theory*, S. Wall and G. Klosko (eds.), Oxford: Rowman & Littlefield: 167-189

Knight, H., 2014, *How Humans Respond to Robots: Building Public Policy through Good Design*,
BROOKINGS REPORT

Kolb, M., 2012, *SOLDIER AND ROBOT INTERACTION IN COMBAT ENVIRONMENTS*,
Dissertation, University of Oklahoma (2012).

Koops, B.-J., Hildebrandt, M., & Jaquet-Chifelle, D.-O., 2010, “Bridging the accountability gap: Rights for new entities in the information society” *Minnesota Journal of Law, Science and Technology*

Krahé, B., Temkin, J. and Bieneck, S., 2007, “Schema-driven information processing in judgements about rape”, *Appl. Cognit. Psychol.* 21: 601-619

Kramer, M. H., 1998, “Rights Without Trimmings”, in Kramer, Simmonds, and Steiner 1998: 7–112

——, 2001, “Getting Rights Right”, in *Rights, Wrongs and Responsibilities*, Matthew H. Kramer (ed.),
London: Palgrave Macmillan UK, 28–95. doi:10.1057/9780230523630_2

——, 2010, “Refining the Interest Theory of Rights”, *The American Journal of Jurisprudence*, 55(1): 31–
39. doi:10.1093/ajj/55.1.31

——, 2013, “Some Doubts about Alternatives to the Interest Theory of Rights”, *Ethics*, 123(2): 245–
263. doi:10.1086/668705

- , 2017, “In Defense of the Interest Theory of Right-Holding: Rejoinders to Leif Wenar on Rights”, in McBride 2017: 49–84 (ch. 3)
- Kramer, M. H., N. E. Simmonds, and H. Steiner, 1998, *A Debate over Rights: Philosophical Enquiries*, Oxford: Clarendon Press
- , 2001, “Getting Rights Right” in M. Kramer (ed.) *Rights, Wrongs and Responsibilities*, Palgrave
- Kurki, V., 2019, *A Theory of Legal Personhood*, Oxford University Press
- Lackey, J., 2006, “Learning from words”, *Philosophy and Phenomenological Research* 73(1): 77–101
- , 2008, *Learning From Words: Testimony as a Source of Knowledge*, Oxford: Oxford University Press
- Lamond, G., 2007, “What is a crime?”, *Oxford Journal of Legal Studies* 27(4): 609–632
- Lancaster, K., 2021, “Non-consensual personified sexbots: an intrinsic wrong”, *Ethics and Information Technology* 23: 589–600
- Langton, R., 1993, “Speech Acts and Unspeakable Acts”, *Philosophy & Public Affairs*, 22(4): 293–330
- , 2009, *Sexual Solipsism: Philosophical Essays on Pornography and Objectification*, Oxford
- Larenz, K., 1983, *Allgemeiner Teil des deutschen Bürgerlichen Rechts. Ein Lehrbuch*, C. H. Beck’sche Verlagsbuchhandlung
- Larmore, C., 1987, *Patterns of Moral Complexity*, Cambridge: Cambridge University Press
- , 1999, “The Moral Basis of Political Liberalism”, *The Journal of Philosophy* Vol. 96, No. 12: 599–625
- Laukyte, M., 2014, “Artificial Agents: Some consequences of a few capacities”, in J. Seibt *et al.* (eds.) *Sociable Robots and the Future of Social Relations*: 115–122, IOS Press

——, 2017, “Artificial agents among us: Should we recognise them as agents proper?”, *Ethics and Information Technology*, 19(1): 1-17

Lawrence v. Texas [2003] 539 U.S. 558

Leland, R.J., 2019, “Civic Friendship, Public Reason,” *Philosophy & Public Affairs*, 47(1): 72–103

Leland, R.J. and H. van Wietmarschen, 2012, “Reasonableness, Intellectual Modesty, and Reciprocity in Political Justification,” *Ethics*, 122(4): 721–747

——, 2017, “Political Liberalism and Political Community,” *Journal of Moral Philosophy*, 14(2): 142–167

Levinas, E., 1989, “Ethics as first philosophy”, in S. Hand (ed.), *The Levinas Reader*, Basil Blackwell, Oxford: 75-87

Levy, D., 2007, *Love and Sex with Robots*, Harper Perennial, New York

Lingis, A., 1994, *The Community of Those Who Have Nothing in Common*, Indiana University Press, Bloomington and Indianapolis

List, C., 2021, “Group Agency and Artificial Intelligence”, *Philosophy & Technology* 34: 1213-1242

List, C. and Pettit, P., 2011, *Group Agency: The possibility, design, and status of corporate agents*, OUP

Lister, A., 2010, “Public Justification and the Limits of State Action,” *Politics, Philosophy, & Economics*, 9(2): 151–175.

——, 2013, *Public Reason and Political Community*, London: Bloomsbury.

——, 2017, “Public Reason and Reciprocity,” *Journal of Political Philosophy*, 25(2): 155–172

- Lloyd, S., 1998, "Toward a Liberal Theory of Sexual Equality", *Journal of Contemporary Legal Issues* 9: 203-224
- Luck, M., 2009, "The Gamer's Dilemma", *Ethics and Information Technology* 11(1)
- Luck, M. and Ellerby, N., 2013, "Has Bartel Resolved the Gamer's Dilemma?", *Ethics and Information Technology* 15(3)
- Macedo, S., 1995, "Homosexuality and the Conservative Mind", *Georgetown Law Journal* 84
- MacCormick, N., 1977, "Rights in Legislation", in *Law, Morality and Society: Essays in Honour of H.L.A. Hart*, P.M.S. Hacker, and Joseph Raz (eds), Oxford: Clarendon Press: 189–209
- , [1988]2007, *Institutions of Law: An Essay in Legal Theory*, OUP: 77-99
- Machen, A., 1911, "Corporate Personality", *Harvard Law Review* 24(253)
- MacKinnon, C., 1987, "Not a Moral Issue" and "Francis Biddle's Sister: Pornography, Civil Rights and Speech," in *Feminism Unmodified*, Cambridge, MA: Harvard University Press
- , 1989, "Rape: On Coercion and Consent", *Toward a Feminist Theory of the State*, Harvard University Press
- , 1993, *Only Words*, Cambridge: Harvard University Press
- , 2006, "Defining Rape Internationally: A Comment on Akayesu", *Columbia Journal of Transnational Law* 44(3): 940-958
- , 2016, "Rape Redefined", *Harvard Law & Policy Review* 10(2): 431-477

MacWilliam, S., 2023, “The Voice of the ‘Sex Robot’: From Peep-Show Bucket to Willing Victim—The Terrorism of Women’s Speech” in K. Richardson and C. Odland (eds.) *Man-Made Women*, Palgrave Macmillan

Masterpiece Cakeshop, Ltd. v. Colorado Civil Rights Commission [2018], 584 US _ (2018)

Matthias, A., 2004, “The responsibility gap: Ascribing responsibility for the actions of learning automata”, *Ethics and Information Technology* 6: 175-183

McAdams, R. H., 2015, *The Expressive Powers of Law: Theories and Limits*, Harvard University Press, Cambridge MA

McArthur, N., 2017, “The Case for Sexbots”, in J. Danaher and N. McArthur (eds.) *Robot Sex: Social and Ethical Implications*, MIT Press, Cambridge

McBride, M., (ed.), 2017, *New Essays on the Nature of Rights*, Oxford/Portland, OR: Hart Publishing

McGlynn, C., 2017, “Rape Trials and Sexual History Evidence: Reforming the Law on Third-Party Evidence”, *The Journal of Criminal Law* 81(5): 367–392

McKinnon, R., 2013, “The supportive reasons norm of assertion”, *American Philosophical Quarterly* 50(2): 121–135

——, 2014, “Stereotype threat and attributional ambiguity for trans women”, *Hypatia* 29(4): 857–872

——, *The Norms of Assertion: Truth, Lies, and Warrant*, Palgrave Macmillan: Houndmills

——, forthcoming, “How to be an optimist about aesthetic testimony”, *Episteme*

Medina, José, 2012, “Hermeneutical Injustice and Polyphonic Contextualisms: Social Silences and Shared Hermeneutical Responsibilities,” *Social Epistemology: A Journal of Knowledge, Culture and Policy*, 26(2): 201–220

——, 2013, *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination*, Studies in Feminist Philosophy, New York

——, 2016. “Ignorance and Racial Insensitivity,” *The Epistemic Dimensions of Ignorance*, R. Peels and M. Blaauw (ed.), Cambridge: Cambridge University Press, 178–201

Modinos v. Cyprus [1993] 16 EHRR 485

Mikkola, M., 2019, *Pornography: A Philosophical Introduction*, New York

Mill, J. S., 2011[2019], *On Liberty*, available online:
<https://www.gutenberg.org/cache/epub/34901/pg34901-images.html>

Miller v R [2010] EWCA Crim 1578

Mills, C., 2017, *Black Rights/White Wrongs: The Critique of Racial Liberalism*, Oxford: Oxford University Press

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L., 2016, “The ethics of algorithms: Mapping the debate”, *Big Data & Society*, 3(2), 1–21. <https://doi.org/10.1177/2053951716679679>

Mori, M., 1970, “The Uncanny Valley”, *Energy*, Vol. 7, No. 4: 33–35 (in Japanese), K. MacDorman and N. Kageki (trans.), 2012, IEEE ROBOTICS & AUTOMATION MAGAZINE: 98-100

Morin, R., 2016, “Can Child Dolls Keep Pedophiles from Offending?”, *The Atlantic*,
<https://www.theatlantic.com/health/archive/2016/01/can-child-dolls-keep-pedophiles-from-offending/423324/>

Morrigan, V., 2023, “Patriarchal Imaginaries Beyond the Human: ‘Sex’ Robots, Fetish and Fantasy in the Domination and Control of Women”, in K. Richardson and C. Odland (eds.) *Man-Made Women*, Palgrave Macmillan

Murray, L., 2017, “NAE WAY A sex robot has been given a SCOTTISH accent to sound less robotic”, *The Scottish Sun*, 29 October 2017

<https://www.thescottishsun.co.uk/news/scottish-news/1766829/sex-robot-scottish-accent-sounds-less-robotic/>

Naffine, N., 2009, *Law's Meaning of Life: Philosophy, Religion, Darwin and the Legal Person*, Hart Publishing

Neufeld, B.E. and C. Van Schoelandt, 2014, “Political Liberalism, Ethos Justice, and Gender Equality,” *Law and Philosophy*, 33(1): 75–104

Neuner, J., 2013, “Zur Rechtsfähigkeit des Anencephalus”, 31 *Medizinrecht*

Nevejans, N., 2016, “European Civil Law Rules in Robotics”, Directorate General for Internal Policies 2016, Study for the JURI Committee PE 571.379

Nonhuman Rights Project v. Lavery [2014], *People ex rel. Nonhuman Rights Project, Inc. v Lavery* [2014] 124 A.D.3d 148, 998 N.Y.S.2d 248

Noonan, J., 1976, *Persons and Masks of the Law*, New York: Farrar, Straus and Giroux

Norris v. Ireland [1988] 6/1987/129/180

NSPCC 2023, “Protecting children from sexual abuse”, <https://learning.nspcc.org.uk/child-abuse-and-neglect/child-sexual-abuse>

Nussbaum, M., 1995, “Objectification”, *Philosophy & Public Affairs* 24

———, 2011, *Creating Capabilities*, Cambridge, MA: Harvard University Press

- Odland, C. and Richardson, K., 2023, "The End of Sex Robots—For the Dignity of Women and Girls", in K. Richardson and C. Odland (eds.) *Man-Made Women: The Sexual Politics of Sex Dolls and Sex Robots*, Palgrave Macmillan
- Oosterheld, C., 2016, "Backup utility functions as a fail-safe AI technique", in *Foundational Research Institute, Report FRI-16-2*
- Okin, S., 1994, "Political Liberalism, Justice, and Gender", *Ethics* 105(1): 23-43
- , 2004, "Justice and Gender: An Unfinished Debate", *Fordham Law Review* 72: 1537-1567
- Olivecrona, K., 1928, *Studier över begreppet juridisk person i romersk och modern rätt*, Appelbergs
- ONS, 2021, "Sexual offences in England and Wales overview: year ending March 2020", Available online:
<https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/sexualoffencesinenglandandwalesoverview/march2020>
- Open Letter to the EU Commission, 2018, "OPEN LETTER TO THE EUROPEAN COMMISSION: ARTIFICIAL INTELLIGENCE AND ROBOTICS", Available online:
<http://www.robotics-openletter.eu/>
- Orts, E., 2013, *Business Persons. A Legal Theory of the Firm*, OUP
- Pallikkathayil, J., 2016, "Revisiting the Interest Theory of Rights: Discussion of The Morality of Freedom", *Jerusalem Review of Legal Studies*, 14(1): 147–157
- Parikh, R., 2002, "Social Software", *Synthese* 13(2): 187-211
- Patridge, S., 2010, "The Incurable Social Meaning of Video Game Imagery", *Ethics and Information Technology* 13(4): 303-312

——, 2013, “Pornography, Ethics and Video Games”, *Ethics and Information Technology* 15(1)L 25-34

Patridge, S. and Jordan, A., 2010, “Against the moralistic fallacy: In Defence of Modest Moralism about Humour”, *Ethical Theory and Moral Practice* 15(1): 84-95

Peeters, A. and Haselager, P., 2021, “Designing Virtuous Sex Robots”, *International Journal of Social Robotics* 13: 55-66

Quong, J., 2004, “The Scope of Public Reason”, *Political Studies* 52: 233-250

——, 2011, *Liberalism Without Perfection*, Oxford: Oxford University Press.

——, 2012, “*Liberalism Without Perfection*: Replies to Gaus, Colburn, Chan, Bocchiola,” *Philosophy & Public Issues*, 2(1): 51–79.

——, 2013, “On The Idea of Public Reason,” in *The Blackwell Companion to Rawls*, J. Mandle and D. Reidy (eds.), Oxford: Wiley-Blackwell, pp. 265–280.

——, 2014, “What is the Point of Public Reason?” *Philosophical Studies*, 170(3): 545–553

R v Brown [1993] UKHL 19, 1 AC 212

Rawls, J., 1993[2005], *Political Liberalism*, Columbia Classics in Philosophy

——, 1995, “Reply to Habermas,” *The Journal of Philosophy*, 92(3): 132–180

——, 1997, “The Idea of Public Reason Revisited”, *The University of Chicago Law Review*, Vol. 64, No. 3: 765-807

——, 1999a, *A Theory of Justice: Revised Edition*, Oxford: Oxford University Press

——, 1999b, *Collected Papers*, S. Freeman (ed.), Cambridge, MA: Harvard University Press

——, 2001, *Justice as Fairness: A Restatement*, Cambridge, MA: Harvard University Press

Raz, Joseph, 1970, *The Concept of a Legal System: An Introduction to the Theory of Legal System*, Oxford: Clarendon Press

——, 1986, *The Morality of Freedom*, Oxford: Clarendon Press

——, 1994, *Ethics in the Public Domain: Essays in the Morality of Law and Politics*, Oxford: Clarendon Press; reprinted, 2001

——, 1995, *Legal Rights*, in ETHICS IN THE PUBLIC DOMAIN: ESSAYS IN THE MORALITY OF LAW AND POLITICS 259

Realbotix, 2023, <https://realbotix.com/>

Reliable Cons. v. Earle [2008] 517 F.3d 738 (5th Cir. 2008)

Restatement (Third) of Agency, 2006, Available online: <https://studylib.net/doc/8168533/restatement-of-the-law---agency-restatement--third--of-a...>

Ribeiro, M., Singh, S. and Guestrin, C., 2016, “*Why Should I Trust You?*” *Explaining the Predictions of Any Classifier* <https://arxiv.org/pdf/1602.04938.pdf>

Richardson, K., 2015, “The Asymmetrical ‘Relationship’: Parallels Between Prostitution and the Development of Sex Robots”, *SIGCAS Computers & Society* 45(3): 290-293

——, 2016, “Sex robot matters: slavery, the prostituted, and the rights of machines”, *IEEE Technology and Society Magazine* 35(2): 46-53

——, 2023, “The End of Sex Robots: Porn Robots and Representational Technologies of Women and Girls”, in K. Richardson and C. Odland (eds.) *Man-Made Women: The Sexual Politics of Sex Dolls and Sex Robots*, Palgrave Macmillan

Richardson, K., and C. Odland (eds.), 2023, *Man-Made Women: The Sexual Politics of Sex Dolls and Sex Robots*, Palgrave Macmillan

Riehm, T., 2020, “Riehm: Nein zur Person!”, *Recht Digital* 42

Robeyns, I. and Byskov, M., “The Capability Approach”, *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.)

Rochfeld, J., 2013, *Les grandes notions du droit privé*, 2nd ed. Presses Universitaires de France

Roe v Wade [1972] 410 US 113

Roff, H., 2013, “Killing in War: Responsibility, Liability and Lethal Autonomous Robots” in F. Allhoff, N. G. Evans and A. Henschke (eds.) *Routledge Handbook of Ethics and War: Just War Theory in the Twenty-First Century*, New York: Routledge 352-364

Russell, S. J., and Norvig, P., 2009, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall Press

SAE, 2021, “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles”, Society for Automotive Engineers, Available online: https://www.sae.org/standards/content/j3016_202104/

Saerbeck, M., and Bartneck, C., 2010, “Attribution of Affect to Robot Motion”, PROCEEDINGS OF THE 5TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI2010), Osaka: 53-60

Salmond, J. W., 1920, *Jurisprudence*, Sweet and Maxwell

Sandberg, A., 2015, “Sex and death among the robots: When should we campaign to ban robots?”, *Practical Ethics Blog*

Sandel, M., 2012, *What Money Can't Buy*, London: Penguin

Santa Clara Co v. Southern Pacific Railroad [1886], 118 US 394

Santoni de Sio, F., & Mecacci, G., 2021, “Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them”, *Philosophy & Technology*, 34: 1057-1084

Savigny, F. C. von, 1840, *System des heutigen Römischen Rechts*, Veit

——, 1867, *System of the Modern Roman law, Volume 1*, W. Holloway (tr.), J. Higginbotham (ed.)

——, 1884, *Jural Relations; or, the Roman Law of Persons as Subjects of Jural Relations: Being a Translation of the Second Book of Savigny's System of Modern Roman Law*, W. H. Rattigan (tr., ed.), Wildy & Sons

Scanlon, T. M., 1998, *What We Owe to Each Other*, Harvard University Press

Schellekens, M., 2018, “No-fault compensation schemes for self-driving vehicles”, *Law, Innovation and Technology*, 10(2), 314–333. <https://doi.org/10.1080/17579961.2018.1527477>

Scherer, M., 2017, “Is AI Personhood Already Possible Under U.S. LLC Laws? (Part One: New York)”, 14 May 2017

Schmitt, F., 2006, “Testimonial justification and transindividual reasons”, *The Epistemology of Testimony*, J. Lackey and E. Sosa (eds.) Oxford: Oxford University Press: 193–224

Schouten, G., 2013, “Restricting Justice: Political Interventions in the Home and in the Market,” *Philosophy & Public Affairs*, 41(4): 357–388.

——, 2017, “Citizenship, Reciprocity, and the Gendered Division of Labor: A Stability Argument for Gender Egalitarian Interventions,” *Politics, Philosophy, & Economics*, 16(2): 174–209.

——, 2019, *Liberalism, Neutrality, and the Gendered Division of Labor*, New York: Oxford University Press

Sen, A., 1992, 1992, *Inequality Re-examined*, Oxford: Clarendon Press

Sexual Offences Act 2003, Available online:
<https://www.legislation.gov.uk/ukpga/2003/42/part/1/crossheading/rape>

Sharkey, N., van Wynsberghe, A., Robbins, S., and Hancock, E., 2017, “Our sexual future with robots”, The Hague: Foundation for Responsible Robotics

Available online here: <https://www.scribd.com/document/353020582/Report-Our-Sexual-Future-With-Robots>

Shen, F. X., 2019, “Government officials have no idea how to regulate the growing sex-robot industry”, *The Business Insider*, Feb. 12 2019

<https://www.businessinsider.com/sex-robot-industry-regulations-2019-2?r=US&IR=T>

Shor, E. and Seida, K., 2019, “‘Harder and Harder?’ Is Mainstream Pornography Becoming Increasingly Violent and Do Viewers Prefer Violent Content?”, *The Journal of Sex Research* 56: 16–28

———, 2020, *Aggression in Pornography: Myths and Realities*, Routledge

Simmonds, N. E., 1998, “Rights at the Cutting Edge”, in Kramer, Simmonds, and Steiner 1998: 113–232

Sinnott-Armstrong, W. and Skorburg, J. A., 2021, “How AI can aid bioethics”, *Journal of Practical Ethics*

Smith, B., 1928, “Legal Personality”, *The Yale Law Journal* Vol. 37, No. 3: 283-299

Smith, O. and Skinner, T., 2017, “How rape myths are used and challenged in rape and sexual assault trials”, *Social & Legal Studies*

- Solum, L. B., 1992, “Legal Personhood for Artificial Intelligences”, 70 NC L Rev 1231
- Sparrow, R., 2004, “The Turing Triage Test”, *Machine Ethics and Robot Ethics*: 95-105
- , 2007, “Killer Robots”, *Journal of Applied Philosophy*, 24: 62-77
- , 2016, “Robots and Respect: assessing the case against autonomous weapon systems”, *Ethics & International Affairs* 30: 93-116
- , 2017, “Robots, rape, and representation”, *International Journal of Social Robotics* 9(4): 465-477
- , 2019, “Regulating Child Sex Robots: Restriction or Experimentation?”, *Med Law Rev* 27(4): 553-575
- Srinivasan, A., 2021, *The Right to Sex: Feminism in the Twenty-First Century*, New York: Farrar, Straus and Giroux
- State v Horton* [1905] 51 S.E. 945
- Steiner, H., 1994, *An Essay on Rights*, Oxford/Cambridge, MA: Blackwell
- Stemplowska, Z., 2020, “Duties to the Dead: Is Posthumous Mitigation of Injustice Possible?”, *Oxford Studies in Political Philosophy* 6, Oxford University Press
- Sterri, A. and Earp, B., 2021, “The ethics of sex robots”, in *The Oxford Handbook of Digital Ethics*, Oxford: Oxford University Press
- Stevens, R., 2007, *Torts and Rights*, Oxford: Oxford University Press
- Stilgoe, J., 2020, “Who Killed Elaine Herzberg?” In *Who’s Driving Innovation?* (pp. 1–6). Springer International Publishing. https://doi.org/10.1007/978-3-030-32320-2_1

Stone, C., 1972[2010], “Should Trees Have Standing?”, in *Should Trees Have Standing? Law, Morality, and the Environment*, 3rd edition, OUP

Strikwerda, L., 2017, “Legal and Moral Implications of Child Sex Robots”, in J. Danaher and N. McArthur (eds.) *Robot Sex: Social and Ethical Implications*, MIT Press, Cambridge

Green, S., 2020, “Legal Moralism, Overinclusive Offenses, and the Problem of Wrongfulness Conflation”, *Criminal Law & Philosophy* 14: 417-430

Sturges, F., 2020, “Sex Robots & Vegan Meat by Jenny Kleeman review – the future of food, birth and death?”, *The Guardian*, 1 July 2020

<https://www.theguardian.com/books/2020/jul/01/sex-robots-vegan-meat-by-jenny-kleeman-review-the-future-of-food-birth-and-death>

Sumner, L., 1987, *The Moral Foundations of Rights*, Oxford: Clarendon Press.

Sun, C., Bridges, A., Johnson, J. A. and Ezzell, M. B., 2016, “Pornography and the Male Sexual Script: An Analysis of Consumption and Sexual Relations”, *Archives of Sexual Behavior*, 45: 983–94

Sunstein, C., 1996, “On the Expressive Function of Law”, *University of Pennsylvania Law Review*, Vol. 144, No. 5: 2021-2053

Tadros, V., 2016, *Wrongs and Crimes*, OUP

Tan, S. H., 2002, “Validity and Obligation in Natural Law Theory”, *Regent University Law Review* 15: 195-221

Timmins, B., 2017, “New sex robots with “Frigid” settings allow men to simulate rape”, *The Independent* 19 July, <https://www.independent.co.uk/life-style/sex-robots-frigid-settings-rape-simulation-men-sexual-assault-a7847296.html>

Tollefsen, D., 2002, “Collective intentionality and the social sciences”, *Philosophy of the Social Sciences* 32(1): 25-50

——, 2015, *Groups as agents*, Polity Press

Tonkens, R., 2012, “Out of character: on the creation of virtuous machines”, *Ethics of Information Technology* 14(2): 137-149

Toonen v. Australia [1994] Communication No. 488/1992, 4 April 1994: Australia

Tracy, P., 2017, “Controversial sex robot that can orgasm is now for sale”, *Daily Dot*

<https://www.dailydot.com/debug/sex-robot-samantha-on-sale/>

Trudy, J., Brown, R., Ibrahim, I. and Parellada, O., 2021, “A Sandbox Approach to Regulating High-Risk Artificial Intelligence Applications”, *European Journal of Risk Regulation*

Turing, A., 1950, “Computing machinery and intelligence”, *Mind* 59(236): 433-460

Turkle, S., 2010, “In Good Company?: On the Threshold of Robotic Companions”, in *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*, Y. Wilks (ed.), John Benjamins Publishing Company

Turner, J., 2019, *Robot Rules: Regulating Artificial Intelligence*, Palgrave Macmillan

Tnyman v. Tnyman [1993] 855 S.W.2d 619

UK Parliament Home Affairs Committee, 2022, “Investigation and prosecution of rape”, *Eighth Report of Session 2021-22*, Available online: <https://committees.parliament.uk/publications/9600/documents/166175/default/>

UN Libya Report 2021, *Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council*: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/037/72/PDF/N2103772.pdf?OpenElement>

Upton, J., Hazell, A., Abbott, R., and Pilling K., 2021, “Independent report: The relationship between pornography use and harmful sexual behaviours”, Available online:

<https://www.gov.uk/government/publications/research-publications/the-relationship-between-pornography-use-and-harmful-sexual-behaviours#the-relationship-between-pornography-use-and-harmful-sexual-behaviours>

US Department of Defense DoD 2012, *Directive Number 3000.09 on Autonomy in Weapon Systems*, November 21 2012

<https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf>

Vadas, M., 2005, “The Manufacture-for-Use of Pornography and Women’s Inequality”, *Journal of Political Philosophy* 13: 174-193

Vallentyne, P., 2007, “Libertarianism and the state”, *Social Philosophy and Policy*, 24: 187–205

Vallier, K., 2011, “Against Public Reason Liberalism’s Accessibility Requirement,” *Journal of Moral Philosophy*, 8(3): 366–389.

—, 2014, *Liberalism and Public Faith: Beyond Separation*, New York: Routledge.

—, 2016, “Public Justification vs Public Deliberation: The Case for Divorce,” *Canadian Journal of Philosophy*, 45(2): 139–158.

—, 2017, “On Jonathan Quong’s Sectarian Political Liberalism,” *Criminal Law and Philosophy*, 11(1): 175–194.

—, 2019, *Must Politics Be War? Restoring Our Trust in the Open Society*, New York: Oxford University Press

Vallor, S., 2016, “Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting”, Oxford University Press, Oxford

Vera-Gray, F., McGlynn, C., Kureshi, I., and Butterby, K., 2021, “Sexual violence as a sexual script in mainstream online pornography”, *The British Journal of Criminology* 61: 1243-1260

Vienna Convention on the Law of Treaties, 1969, Available online: https://legal.un.org/ilc/texts/instruments/english/conventions/1_1_1969.pdf

WAFEW, 2021, “Written evidence submitted by Women’s Aid Federation of England” (INV0010), Available online: <https://committees.parliament.uk/writtenevidence/36735/pdf/>

Wagner, G., 2019, “Robot, Inc.: Personhood for Autonomous Systems?”, 88 *Fordham L Rev* 591

Wall, S., 2013a, “Enforcing Morality”, *Criminal Law and Philosophy* 7: 455-471

—, 2013b, “Moral Environmentalism” in C. Coons and M. Weber, *Paternalism: In Theory and Practice*, Cambridge: Cambridge University Press

—, 2021, “Perfectionism in Moral and Political Philosophy,” *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), E. N. Zalta (ed.)

<https://plato.stanford.edu/archives/fall2021/entries/perfectionism-moral>

Wallach, W. and Allen, C., 2009, *Moral Machines: Teaching Robots Right from Wrong*, Oxford University Press, Oxford

Watson, L., and C. Hartley, 2018, *Equal Citizenship and Public Reason: A Feminist Political Liberalism*, New York: Oxford University Press

- Wellman, C., 1985, *A Theory of Rights: Persons under Laws, Institutions, and Morals*, Totowa, NJ: Rowman & Allanheld
- , 1995, *Real Rights*, New York: Oxford University Press
- Wellman, C., 2014, “Procedural Rights”, *Legal Theory*, 20(4): 286–306
- Wenar, L., 2005, “The Nature of Rights”, *Philosophy & Public Affairs*, 33(3): 223–252
- , 2005[2021], “Rights”, *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), E. N. Zalta (ed.)
- West, C., 2022, "Pornography and Censorship", *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), E. N. Zalta and U. Nodelman (eds.)
- Wise, S. M., 1998, “Hardly a Revolution—The Eligibility of Nonhuman Animals for Dignity-Rights in a Liberal Democracy”, *Vermont Law Review* 22: 793
- , 2010, “Legal Personhood and the Nonhuman Rights Project”, *Animal Law Review* 17(1)
- Wood, T., Deep AI: <https://deepai.org/machine-learning-glossary-and-terms/unsupervised-learning>
- Wright, P. J., Tokunaga, R. S., Kraus, A., 2016, “A meta-analysis of pornography consumption and actual acts of sexual aggression in general population studies”, *J. Commun* 66(1): 183-205
- Wyre, R., 1992, “Pornography and Sexual Violence: Working with Sex Offenders”, in C. Itzin (ed.) *Pornography: Women, Violence and Civil Liberties*, Oxford: Oxford University Press
- Yuracko, K., 2003, *Perfectionism and Contemporary Feminist Values*, Bloomington: Indiana University Press
- Zelizer, V., 1996, *The Social Meaning of Money*, Princeton: Princeton University Press

Zhang, E., 2021, “Why Sex? Sex-Bots from a Daoist Perspective”, in R. Fan and M. Cherry (eds), *Sex Robots: Social Impact and the Future of Human Relations*, Springer

Zhou, Y. and Paul, B., 2016, “Lotus Blossom or Dragon Lady: A Content Analysis of ‘Asian Women’ Online Pornography”, *Sexuality and Culture*, 20: 1083–100