

THE SCOPE OF JUSTICE:
WHOM SHOULD RIGHTS PROTECT?

by

MATTHEW TAYLOR

A thesis submitted to the
Department of Philosophy
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
September 2015

Copyright © Matthew Taylor, 2015

Abstract

This thesis argues that the strongest account of moral rights entails that animals and other marginal cases hold rights. The thesis contends that mutual advantage social contract theories offer the strongest account of rights from a security perspective, and that such theories entail rights for animals and marginal cases. Both of these claims are widely contested. Chapter 1 examines the fundamental elements of a social contract theory as developed by Hobbes and Hume. Chapter 2 revises the fundamental elements of contract theory to make them more persuasive, and derives from them an account of rights for animals and marginal cases. Chapter 3 examines the most plausible competing accounts of rights for animals and marginal cases: utilitarianism, neo-Kantianism, and capability theory. Chapter 4 argues that the mutual advantage account of rights is better than the leading competitors from a security perspective.

Acknowledgments

Research for this thesis was supported by the Social Sciences and Humanities Research Council.

I would like to thank Will Kymlicka, Christine Sypnowich, and Alistair Macleod for their comments, guidance, and support in the writing of this thesis. I would also like to thank Susan Dimock and Jeffrey Collins for their comments.

Contents

| | |
|---|------------|
| Abstract | i |
| Acknowledgments | ii |
| Contents | iii |
| Chapter 1: Introduction | 1 |
| 1.1 Moral Rights | 2 |
| 1.2 Marginal Cases | 6 |
| 1.3 Mutual Advantage Social Contract Theory | 8 |
| 1.4 Security and Non-Ideal Theory | 11 |
| Chapter 2: Introduction to Contractarianism: Hobbes and Hume | 18 |
| 2.1 Hobbes | 19 |
| 2.1.1 Why Begin With a State of Nature? The Resolutive-Compositive Method | 19 |
| 2.1.2 Human Nature | 21 |
| 2.1.3 The Problem of the State of Nature: War | 30 |
| 2.1.4 The Right of Nature | 34 |
| 2.1.5 The Laws of Nature | 40 |
| 2.1.6 Contract and Cooperation in the State of Nature | 44 |
| 2.1.7 The Source of All Constraint: Mutual Advantage | 52 |
| 2.2 Hume | 57 |
| 2.2.1 Circumstances of Justice | 58 |
| 2.2.2 Convention | 66 |
| 2.3 Those Excluded From Moral Rights | 68 |
| 2.4 Conclusion | 75 |
| Chapter 3: The Contractarian Theory of Animal Rights | 76 |
| 3.1 Contractarianism Revisited | 77 |
| 3.1.1 Human Nature and Rights | 77 |
| 3.1.2 Human Nature, Misanthropic Desires, and Coalition Games | 82 |

| | | |
|---|--|------------|
| 3.1.3 | Why Consider a State of Nature? | 91 |
| 3.1.4 | The Revised Contractarian Theory of Moral Rights | 93 |
| 3.1.5 | Mutual Advantage Contractarianism and Rights | 96 |
| 3.1.6 | Rights v. Protectorate Status | 100 |
| 3.1.7 | Protectorate Status | 101 |
| 3.1.8 | Summary | 109 |
| 3.2 | Direct Moral Rights for Animals | 110 |
| 3.2.1 | Benefits | 112 |
| 3.2.2 | Species | 121 |
| 3.2.3 | Summary | 127 |
| 3.3 | Protectorate Status | 128 |
| 3.4 | Objections | 132 |
| 3.4.1 | Argument from Marginal Cases | 132 |
| 3.4.2 | Conflation | 135 |
| 3.5 | Conclusion | 137 |
| Chapter 4: Dominant Theories of Animal Rights | | 139 |
| 4.1 | Utilitarian Theory of Rights | 139 |
| 4.2 | Kantian Theory of Rights | 145 |
| 4.3 | Capabilities Theory of Rights | 149 |
| 4.4 | Conclusion | 160 |
| Chapter 5: Evaluating Competing Theories of Rights | | 161 |
| 5.1 | Desiderata For Theories of Rights | 161 |
| 5.1.1 | Security: The Central Desideratum | 161 |
| 5.1.2 | How Do “Protections” Fit In? | 167 |
| 5.1.3 | Criteria For Evaluating Elements of Theories of Rights | 167 |
| 5.2 | Evaluating Competing Theories | 173 |
| 5.2.1 | Utilitarianism | 175 |
| 5.2.2 | Kantianism | 192 |
| 5.2.3 | Capabilities Theory | 197 |
| 5.2.4 | Contractarianism | 205 |
| 5.2.5 | Final Evaluation | 211 |
| Chapter 6: Conclusion | | 216 |
| Bibliography | | 223 |

Chapter 1

Introduction

This thesis starts with a simple question: “What moral obligations, if any, do humans have toward animals?” Answering this question quickly turns into a complicated affair.

This thesis aims to offer the most plausible account of how we ought to treat animals, humans with disabilities, and generally all of the other “marginal cases.”¹ The best way to support such an ethical account, I contend, is to ground it in mutual advantage contractarianism. Starting from a mutual advantage contractarian framework, this thesis will develop an account of the rights and protections belonging to animals and other marginal cases.

This thesis sets itself two main goals. The first goal is to develop an account of marginal rights grounded in mutual advantage contractarianism. The second goal is to argue that this account of marginal rights is better than all competing accounts. For the sake of the first goal, chapter 2 examines the fundamental elements of mutual advantage contractarianism as developed by Hobbes and Hume, and from these

1. The term “marginal cases” is given a preliminary definition in section 1.2 on page 6, and discussed more thoroughly in section 3.4.1 on page 132.

elements chapter 3 derives our moral obligations toward animals and marginal cases. For the sake of the second goal, chapter 4 examines the competing accounts of rights for animals and marginal groups, and chapter 5 argues that mutual advantage contractarianism offers the better account of rights from a security perspective.²

In this introduction, I hope to clarify a number of concepts and assumptions to which I will appeal in subsequent chapters.

1.1 Moral Rights

This thesis examines the moral rights of animals and marginal cases. It is important to specify what is meant by “moral” and “rights.”

There are three constraints governing the definitions I offer here: fit with mutual advantage contractarianism, fit with the other theories discussed here, and relevance to real-world practices. Firstly, the kind of social contract argument developed in this thesis leads to a certain way of thinking about moral rights. My definitions should fit with the contractarian theory, and should help make clear to the reader what the theory means when it talks about moral rights.

However, as the second constraint notes, it is important that the definitions should not be tailored so narrowly to contractarianism that they exclude the alternative theories discussed here by definition. The thesis wants to compare the contractarian account of moral rights to alternative accounts, and to argue that the contractarian account is best. In order for this comparison to be possible, the theories discussed must be talking about the same thing. Otherwise, the theories are talking past one another and comparison is impossible.

2. The term “security” is defined in section 5.1.1 on page 161.

Thirdly, the argument advanced by this thesis aims to be relevant to the way we behave in the real world. It argues that we ought to respect certain duties toward animals and marginal cases. While I do not think that common usage has any normative significance, it is important that my definitions relate to real-world practices. Thus, while I am not trying to offer definitions that capture all and only what we typically mean when we use the terms in common language, the definitions should not be so unrelated to common usage that my theory becomes irrelevant to the real world.

These three constraints lead to a rather sparse preliminary definition of moral, but bear in mind that each theory, including contractarianism, adds further specifications to the definition. According to the preliminary definition of *moral*, rights are moral if and only if two requirements are met. First, the duties stemming from the right in question must be grounded in reason, by which I mean that, in principle, a sound reasoner could be convinced that he or she ought to fulfill the duties in question.³ Second, the right-holder must be made better off by the purported duty-holder fulfilling the duty in question, and must not be the same person as the duty-bearer.

The first requirement rules out pure force as the source of moral obligations. Should I threaten to kill you unless you kiss a frog, any compulsion you faced to kiss the frog would not count as moral. This claim will be subject to a crucial limit, however: there are some individuals with whom cooperation on terms that everyone can accept will be impossible, and thus in order to get any moral rules, they will need to be coerced. Thus, moral rules must be based on reason that is acceptable to all those with whom cooperation on mutually acceptable terms is possible.⁴ This

3. Much more will be said about this first requirement in section 5.1.1 on page 161.

4. See section 3.1.2 for more on this matter.

requirement meets the three constraints on the definitions.

The second requirement narrows the scope of “moral” so that it no longer comprises all that we might want to talk about in ethics. The moral claims that are made in this thesis refer to what has been called “social morality;” they articulate a set of rules that are needed for the sake of peaceful social interaction.⁵ Social morality essentially focuses on our obligations to others—on the actions that we perform that make others better or worse off. This fits well with mutual advantage contractarianism, allowing it to distinguish moral requirements like “do not kill others” from non-social prudential prescriptions, like “accumulate as much money as you can.”⁶ But this restriction to social morality may not fit so well with some of the other theories discussed here—in particular, with Nussbaum’s capability theory which could be seen to be more connected with an individual’s flourishing whether or not others have an interest in that individual’s flourishing. I contend that this restriction is justified, however, by the third constraint on the definitions. In particular, insofar as we are discussing *rights* in particular, as opposed to other moral concepts, we are explicitly talking about social morality. Nussbaum is focused on articulating a theory of rights—and indeed, an explicitly *political* theory of rights—and thus her argument is appropriately interpreted as one concerning social morality. This focus on social morality qualifies all of the conclusions reached in this thesis: the thesis does not claim to discuss all of the morally relevant aspects of our interactions with animals and marginal cases. It claims only to discuss implications of social morality on our

5. Gerald Gaus, *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (Cambridge: Cambridge University Press, 2012), 2–14. My explanation of “social morality” may differ from Gaus’s—in particular, Gaus means something different by “moral.”

6. Both of these prescriptions can be equally compelling to the actor, but a contractarian can maintain that only the former counts as moral because only in the former case does anyone else have an interest in compelling the actor to comply, and punishing the actor should the actor fail to comply.

interactions with animals and marginal cases.

Thus the preliminary definition of what this thesis means when it talks about “morality” is that it is talking about rights and duties grounded in reason in the domain of social morality. Now it is important to offer a preliminary definition of “rights.”

Firstly, when this thesis talks about rights, it is talking about *basic* rights in the sense that it is talking about rights that do not arise from particular roles or actual contracts. If you and I sign a contract that tomorrow I will give you twenty dollars and you will give me a book, then there is a sense in which you have a right to my twenty dollars tomorrow and I have a right to the book. But these are not the kinds of rights that this thesis concerns. This may appear odd, especially since the thesis is articulating a *contract* account of rights. Section 2.2.2 will discuss Hume’s concept of a convention and how it differs from promises and contracts more generally.⁷ The rights discussed in this thesis also differ from rights grounded in roles or status, like the prime minister’s right to declare war for instance. The rights discussed here are not rights that right-holders hold in virtue of some action they have performed or some role that they occupy. The rights that this thesis discusses are *basic* in the sense that they are rights that right-holders hold only by virtue of being right-holders—the kind of thing entitled to rights. Each theory identifies different properties that entitle one to hold rights.

The rights that this thesis discusses are also basic in the sense that they do not rely on political institutions for their normativity. They are *moral* rights; their force would hold in the absence of a state. By talking about basic rights, I do not mean to

7. Page 66. See also section 2.1.5 on page 40, and section 2.1.7 on page 52, and section 3.1.4 on page 93.

imply that there are no non-basic rights. This thesis is focused on basic rights, but this does not rule out other non-basic rights like citizenship rights or rights derived from specific contracts.

I also assume that rights and duties are correlated: a right is useless unless it generates duties with which others must comply. And all *basic* moral duties—those that do not arise from particular contracts or organizational roles, that hold independently of political institutions, and that are owed to someone other than the duty-bearer him- or herself—correlate with basic rights. Insofar as this thesis focuses exclusively on the basic moral rights of social morality, these assumptions should not be problematic, though I concede that they may not hold beyond this limited domain.

Finally, this thesis does not address third-party obligations, thus the claims made here about basic moral rights do not directly entail any form of enforcement. While I may have a basic moral right against theft, this does not directly entail that, should you steal something of mine, other people are obligated to punish you or anything else. While I think that the claims about basic moral rights have a strong bearing on questions of third-party obligations, I think that a further argument is always needed to justify third-party obligations, and I do not develop such arguments here.

1.2 Marginal Cases

This thesis examines the moral rights of animals and “marginal cases.” This term comes from what is known as “the argument from marginal cases,” which targets theories that try to restrict rights to all and only humans. The argument, which will be discussed in section 3.4.1,⁸ essentially binds the moral fates of animals and

8. Page 132.

various groups of humans who are referred to as “marginal cases,” typically including human infants, severely disabled humans, ill humans, and elderly humans. Theorists who try to guarantee rights for all and only humans do so by citing some property—typically some form of rationality—and claiming that possession of this property is necessary and sufficient for a being to be owed moral consideration. The argument from marginal cases then claims that whatever property we cite as the basis for possessing rights, either all humans and some animals will possess it, or else all animals and all (or many) human marginal cases will lack it, or else it is an arbitrary property like “humanity.” The argument from marginal cases leads to the conclusion that either some animals and all (or most) marginal cases get rights, or else no animals and no (or few) marginal cases get rights.⁹ The argument from marginal cases is the primary reason why this thesis addresses rights for marginal cases instead of focusing exclusively on animal rights.

Throughout the thesis, I will use “marginal cases” as a shorthand to designate human infants, ill or infirm humans, humans with disabilities, etc—basically, all of the human cases whose membership in the set of beings who receive moral protections is in question. I recognize that the term ‘marginal’ has implications that are less than ideal, suggesting that perhaps *most* people are not like that, and that only a few somewhat negligible individuals fall into this class. I do not intend any of these implications: many have rightly noted that all humans spend time in this category.¹⁰ I will use the word marginal in the sense of falling near the edge of moral consideration, at the limits of the circumstances of justice.

9. Sue Donaldson and Will Kymlicka, *Zoopolis: A Political Theory of Animal Rights* (Oxford: Oxford University Press, 2011), 26-27; Peter Singer, *Animal Liberation* (New York: Ecco, 2002), 16, 18-20.

10. Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 26.

It is important to note that many philosophers who deploy the argument from marginal cases would not endorse the claim that human infants, ill or infirm humans, or humans with disabilities are near the edge of moral consideration. The argument is used to show that endorsing certain demanding properties like rationality as necessary for moral consideration leads to unpalatable conclusions when applied consistently, like the exclusion or marginalization of human infants, ill humans, and humans with disabilities. These philosophers are not endorsing the marginalization of these human cases; they are highlighting this marginalization as a reason to reject such demanding requirements for moral status.

I will focus primarily on animal rights in this thesis, dedicating somewhat less attention to the details of rights for marginal cases, because I find that marginal cases have been better addressed in the existing mutual advantage contractarian literature and that I have little to add.¹¹

1.3 Mutual Advantage Social Contract Theory

This thesis focuses on developing and defending an account of animal rights and rights for marginal cases grounded in mutual advantage social contract theory. It is important to distinguish mutual advantage contract theory from other forms of social contract theory. As Barry and Kymlicka have noted, the fact that all social contract theorists employ the notion of a contract is rather superficial, and distracts us from crucial differences between those contract theories that fall in the category of “Justice as Mutual Advantage,” on the one hand, and those that fall in the category of “Justice

11. See, for instance, Chapter 15, “Children and Rights” in Jan Narveson, *Respecting Persons in Theory and Practice: Essays on Moral and Political Philosophy* (Lanham: Rowman & Littlefield, 2002), 265–278; and Chapter 19, “The Problem of Children” in Jan Narveson, *The Libertarian Idea* (Peterborough, Ontario: Broadview Press, 2001), 269–274.

as Impartiality,” on the other.¹² Theories of justice as mutual advantage ground moral claims in classical rational choice theories, insisting that it must be in our interest—must maximize our individual expected utility—to obey moral principles, and allowing differences in power to impact moral obligations.¹³ Theorists in this camp include, for instance, Hobbes, Hume in part,¹⁴ and Gauthier.¹⁵ Theories of justice as impartiality, on the other hand, appeal to a special moral motive apart from our desire to maximize our expected utility, and insist that moral principles be agreed to without appeal to morally irrelevant factors like differences in power.¹⁶ Theorists in this camp include Hume in part, Rawls, Scanlon, and Barry.¹⁷

This thesis attempts to ground animal rights and rights for marginal cases in mutual advantage contract theory, which is rather rare. Most philosophers who attempt to ground animal rights in contract theory explicitly focus on impartiality contract

12. Brian Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice* (Los Angeles: University of California Press, 1989), 6–8, 269–270; Will Kymlicka, *Contemporary Political Philosophy: An Introduction*, 2nd ed. (Oxford: Oxford University Press, 2002), 68–69, 128–129.

13. Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice*, 6–8.

14. Barry notes that Hume seems to have developed two or three distinct theories of justice, offering a mutual advantage theory with his “circumstances of justice,” an impartiality theory with his “judicious spectator,” and a utilitarian theory built out of the judicious spectator. (ibid., 146.)

15. Thomas Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, ed. Edwin Curley (Indianapolis: Hackett Publishing Company, 1994); David Hume, *A Treatise of Human Nature*, ed. Ernest C. Mossner (London, UK: Penguin Books, 1969); David Hume, *An Inquiry Concerning the Principles of Morals*, ed. Charles W. Hendel (New York: The Liberal Arts Press, 1957); David P. Gauthier, *Morals by Agreement* (Oxford: Clarendon Press, Oxford University Press, 1986).

16. Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice*, 6–8.

17. Hume, *A Treatise of Human Nature*; Hume, *An Inquiry Concerning the Principles of Morals*; John Rawls, *A Theory of Justice*, Revised Edition (Cambridge, Mass.: The Belknap Press of Harvard University Press, 1999); Thomas M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Belknap Press of Harvard University Press, 1998); Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice*.

theories, most notably Rawls's.¹⁸ And those who attempt to ground animal rights in contract theory without marking the difference between justice as mutual advantage and justice as impartiality tend to draw heavily on the impartiality contract theorists, preventing their arguments from applying to mutual advantage theories.¹⁹ Of those few who address the question of rights for animals and marginal cases from a mutual advantage contract perspective, most explicitly exclude animals from qualifying directly for rights, sometimes relegating them to a kind of indirect moral standing that

18. See, for instance, Mark Rowlands, "Contractarianism and Animal Rights," *Journal of Applied Philosophy* 14, no. 3 (1997): 235–247; Julia Driver, "A Humean Account of the Status and Character of Animals," in *The Oxford Handbook of Animal Ethics*, ed. Tom L. Beauchamp and R.G. Frey (Oxford: Oxford University Press, 2011), 145–172; Julie Hilden, "A Contractarian View of Animal Rights: Insuring Against the Possibility of Being a Non-Human Animal," *Animal Law* 14, no. 5 (2007): 5–28; Rebekah Humphreys, "Contractarianism: On the Incoherence of the Exclusion of Non-Human Beings," *Percipi* 2 (2008): 28–38; Paola Cavalieri and Will Kymlicka, "Expanding the Social Contract," *Etica & Animali* 8 (1996): 5–33; Mark Coeckelbergh, "Distributive Justice and Co-Operation in a World of Humans and Non-Humans: A Contractarian Argument for Drawing Non-Humans Into the Sphere of Justice," *Res Publica* 15 (2009): 67–84.

19. See, for instance, Mark Bernstein, "Contractualism and Animals," *Philosophical Studies* 86 (1997): 49–72; Kimberly K. Smith, "Animals and the Social Contract: A Reply to Nussbaum," *Environmental Ethics* 30 (2008): 195–207; Kristin Armstrong Oma, "Between Trust and Dominion: Social Contracts Between Humans and Animals," *World Archaeology* 42, no. 2 (2010): 175–187; Catherine Larrère and Raphaël Larrère, "Animal Rearing as a Contract?," *Journal of Agriculture and Environmental Ethics* 12, no. 1 (2000): 51–58; Raymond Anthony Vonne Lund and Helena Röcklinsberg, "The Ethical Contract As a Tool in Organic Animal Husbandry," *Journal of Agriculture and Environmental Ethics* 17 (2004): 23–49.

this thesis will call “protectorate status.”²⁰ While this thesis will argue that many animals qualify for protectorate status, the protectorate status employed in this thesis is slightly different from the indirect moral consideration others discuss, and this thesis goes further to argue that animals also qualify directly for rights.²¹

1.4 Security and Non-Ideal Theory

Chapter 5 will argue that mutual advantage social contract theory offers the best account of basic moral rights. That requires a different argument from the social contract theory developed in chapters 2 and 3. Chapters 2 and 3 present the contractarian argument that we should obey certain moral principles, but what is needed in chapter 5 is an argument to show that the argument from chapters 2 and 3 is

20. Protectorate status is distinguished from rights in section 3.1.6 and defined in section 3.1.7. For mutual advantage views excluding animals from rights, see Jan Narveson, “Animal Rights,” *Canadian Journal of Philosophy* 7, no. 1 (1977): 161–178; Jan Narveson, “Animal Rights Revisited,” in *Ethics and Animals*, ed. William H. Williams and Harlan B. Miller (Clifton, N.J.: Humana Press, 1983), 45–59; Jan Narveson, “On a Case for Animal Rights,” *The Monist* 70, no. 1 (1987): 31–48. For mutual advantage views relegating animals to something like protectorate status (but different in ways to be examined in section 3.1.7), see Christopher W. Morris, “Moral standing and rational-choice contractarianism,” in *Contractarianism and Rational Choice: Essays on David Gauthier’s Morals By Agreement*, ed. Peter Vallentyne (Cambridge: Cambridge University Press, 1991), 76–95; Christopher W. Morris, “The Idea of Moral Standing,” chap. 9 in *The Oxford Handbook of Animal Ethics*, ed. Tom L. Beauchamp and R.G. Frey (Oxford University Press, 2011), 255–275; Andrew I. Cohen, “Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals,” *Journal of Applied Philosophy* 24, no. 2 (2007): 188–201; Andrew I. Cohen, “Contractarianism and Interspecies Welfare Conflicts,” *Social Philosophy and Policy* 26, no. 1 (2008): 227–257. Finally, Tanner endorses the view that mutual advantage contracts relegate animals to protectorate status and critiques such theories for this failing in Julia Tanner, “Contractarianism and Secondary Direct Moral Standing for Marginal Human and Animals,” *Res Publica* 19 (2013): 141–156.

21. I have only encountered an argument that animals qualify directly for rights in a mutual advantage framework once in C. Tucker and C. MacDonald, “Beastly Contractarianism? A Contractarian Analysis of the Possibility of Animal Rights,” *Essays in Philosophy* 5, no. 2. Animal Ethics, Article 31 (2004). Many of the details of their argument agree with the details of the argument I articulate in chapter 3, though I only found their article after writing chapter 3. Chapter 3 develops the argument in much greater detail, however, and my thesis defends the mutual advantage contract theory of animal and marginal rights, which they do not do in that article.

better than all of the alternative arguments considered in chapter 4. Such an argument must have some perspective, or some framework to which it appeals: in virtue of what considerations is the contractarian account the best? Why are those the relevant considerations? The framework or perspective appealed to in this thesis is explicitly one of non-ideal theory. Rawls defines “ideal theory” as the investigation of what would be just in a “well-ordered society” characterized by “strict compliance” to the rules of justice.²² By contrast, this thesis does not assume compliance, and indeed, sees convincing actors to comply to be one of the most important tasks of a moral theory. The thesis evaluates ethical theories from a practical perspective. From this perspective, the function of an ethical theory is considered to be to secure compliance—to convince people to behave.

This perspective on ethical theory is certainly not the only perspective one could take, nor is it neutral or uncontroversial. Among other ways of evaluating moral theories, one that stands out is a truth-sensitive account: one could think that the purpose of moral theories was to reveal irreducible, independent moral truths. From such a perspective, where moral truth is divorced from practical function, the ability to convince people is irrelevant: it could be that most people are bad at moral reasoning and cling to false beliefs. From such a perspective, the only thing that matters about a moral theory is its proximity to some truth, regardless of how many people disagree, or of the viability of its securing a broad consensus. This thesis’s conclusions are completely restricted to the practical perspective; none of the conclusions defended in this thesis have any impact on claims advanced in other perspectives. Thus, this theory does not claim that contractarianism offers the “true” moral theory in the sense of true that many moral theorists endorse, which typically involves some

22. Rawls, *A Theory of Justice*, 8.

sort of irreducibility to non-moral facts.

I focus on the practical perspective in this thesis largely because I find the objections that Mackie and Harman have presented to “objective” accounts of ethics to be highly compelling, though they remain controversial and disputed. Combined, I think that Harman’s and Mackie’s arguments offer considerable evidence that “objective” accounts of ethics boil down to naked appeals to unjustified intuitions. Mackie and Harman target moral theories that are “objective” in the sense that they posit moral facts that offer unconditional reasons for action, or categorical imperatives.²³ Firstly, Mackie rightly notes that a linguistic analysis of moral terms will neither support the existence of moral facts nor determine what the moral facts are: “The claim to objectivity, however ingrained in our language and thought, is not self-validating.”²⁴ Mackie then offers two arguments against objective moral theories.

Mackie’s first argument is the argument from relativity: “radical differences between first order moral judgments make it difficult to treat those judgments as apprehensions of objective truths.”²⁵ If such persistent disagreement does not lead us to doubt the existence of objective moral facts, it should at the very least make us doubtful that our beliefs about the moral facts are true, and that appeal to moral facts will offer a fruitful way to resolve disputes. To object that, while there is disagreement at the applied level, there is agreement on abstract and general moral principles which are suitable candidates for objective moral facts, from which we derive different applied judgments depending on the particular circumstances in which we find ourselves, Mackie claims, is to offer a disingenuous explanation of what most

23. John L. Mackie, *Ethics: Inventing Right and Wrong* (New York, NY: Penguin Books, 1977), 29; Gilbert Harman, *The Nature of Morality: An Introduction to Ethics* (New York, NY: Oxford University Press, 1977), 6.

24. Mackie, *Ethics: Inventing Right and Wrong*, 35, 19–20.

25. *ibid.*, 36.

objectivist moral theories are doing:

[P]eople judge that some things are good or right, and others are bad or wrong, not because—or at any rate not only because—they exemplify some general principle for which widespread implicit acceptance could be claimed, but because something about those things arouses certain responses immediately in them, though they would arouse radically and irresolvably different responses in others. ‘Moral sense’ or ‘intuition’ is an initially more plausible description of what supplies many of our basic moral judgments than ‘reason’.²⁶

Thus, moral disagreement should make us question whether objective moral facts exist.

The second objection that Mackie offers is the “queerness” objection: if moral values existed, they would be metaphysically and epistemologically “queer” in the sense of being unlike anything else we know of. Metaphysically, objective moral values, if they existed, would be things that, simply by virtue of their being, demanded certain actions, making them unlike anything else that exists.²⁷ Epistemologically, to be aware of such values and the actions they called for would require an entirely unique way of knowing or perceiving:

[H]ow can [we] be aware of this authoritative prescriptivity, of the truth of these distinctively ethical premises or of the cogency of this distinctively ethical pattern of reasoning, [when] none of our ordinary accounts of sensory perception or introspection or the framing and confirming of

26. Mackie, *Ethics: Inventing Right and Wrong*, 37–38.

27. *ibid.*, 38.

explanatory hypotheses or inference or logical construction or conceptual analysis, or any combination of these, will provide a satisfactory answer; ‘a special sort of intuition’ is a lame answer, but it is one to which the clear-headed objectivist is compelled to resort.²⁸

Given the queerness of objective moral facts, and the radical disagreement over moral facts, Hobbes and Hume seem to offer us a better explanation of our intuitive moral judgments: we call good whatever we like, and evil whatever we dislike, thereby projecting our attitudes towards things onto the things themselves.²⁹ Our tastes, and thus our intuitive moral judgments, are socially cultivated and reinforced, leading people to approve of things because they are customary, familiar, and pleasant.³⁰ Thus our moral judgments merely report our own attitudes, and do not reveal objective moral facts.

Harman also argues in favour of this conclusion, and claims that our intuitive moral judgments do not provide us with evidence of any objective facts. Our intuitive moral observations can be fully explained by psychological facts about our beliefs and desires, and do not require us to posit any objective facts about the world to explain them.³¹ Thus we are left with our intuitions, and we know that, outside of the ethical realm, our pre-theoretical intuitions often prove false, and that our intuitions are prey to a host of cognitive biases. Moreover, we know that our moral intuitions differ—sometimes drastically. So while intuitions may offer good starting points for research and argumentation, in the absence of any other evidence to support them, they should not be trusted. Thus, theories grounded in intuitions or in objective moral facts are

28. Mackie, *Ethics: Inventing Right and Wrong*, 38–39.

29. *ibid.*, 42–43.

30. *ibid.*, 36, 43.

31. Harman, *The Nature of Morality: An Introduction to Ethics*, 6–7.

highly problematic.

I find these arguments compelling, but I acknowledge that they are contentious, and I do not want to review or add to the debate over objective moral facts. Instead, in this thesis I set aside debates over objective moral facts, and focus on assessing arguments from the practical perspective. Regardless of whether or not there are moral facts, it is a separate question whether moral theories that examine or reveal those moral facts are the most effective theories at securing our interests. Indeed, in this thesis I will argue that mutual advantage contract theory—a theory which explicitly renounces any claim to tracking moral intuitions—offers us the best theory of moral rights from a practical perspective.

While I do not claim that the practical perspective on ethical theory is the only game in town, I do argue in chapter 5 that it is an important perspective, and one that is at least partially shared by many moral theorists. From this perspective, moral theory is a tool that each of us can use to pursue our interests.³² This leads to a certain set of standards for appraising moral theories based on their ability to secure our interests. Thus, the choice of this perspective is not neutral. The perspective prefers some theories over others—most notably, I claim that it shows that contractarianism is better than the other theories considered. Moreover, it shares many assumptions and fundamental premises with mutual advantage contract theory. There is, I concede, a risk of circularity, but a risk that can scarcely be avoided: the argument from the practical perspective cannot appeal to premises that are contradicted or otherwise ruled out by the contractarian theory itself—otherwise, the argument would be contradictory, hypocritical, and would likely be accused of being disingenuous. So the two arguments cannot help but be closely connected given the

32. The sense of “interest” used in this thesis is limited. See section 2.1.2 for more on interests.

requirement of non-contradiction. I hope that the circularity is not vicious, and that the argument from the practical perspective can at least clarify the relative strengths and weaknesses of the competing theories within that perspective.

Chapter 2

Introduction to Contractarianism: Hobbes and Hume

The purpose of this chapter is to introduce mutual advantage social contract theory so as, first, to identify what it is about these theories that makes them compelling, and second, to understand why some philosophers have taken them to exclude animals, disabled humans, and infants from having moral rights. Because Hobbes and Hume are both prime examples of the kind of theory of justice I am interested in, and primary targets of critics of that theory, Hobbes and Hume present a natural starting point for this thesis.

This thesis is meant primarily to be an investigation of how best to conceive of rights, and of this conception's bearing on who can be a right-holder. I investigate the theories of Hobbes and Hume toward this end, and as a result, my analysis may be less detailed than other primarily historical analyses.

2.1 Hobbes

Hobbes’s *Leviathan* offers the first account of justice as mutual advantage to be developed in any considerable length.¹ Put naively, Hobbes considers individuals in a hypothetical pre-social state of nature, and claims that in such a state, cooperation is nearly impossible and life is “nasty, brutish, and short” (I.xiii.9).² Individuals are strongly motivated to exit the state of nature by forming a government—so strongly motivated that Hobbes thinks that they would find it advantageous to form a totalitarian monarchy with a single sovereign whose power is unrestrained. Of course, what contemporary scholars find attractive about *Leviathan* is not its argument for an all-powerful sovereign; rather, most contemporary scholars think that Hobbes’s argument for such a sovereign is flawed.³ What is of primary interest to this thesis is Hobbes’s moral theory.

2.1.1 Why Begin With a State of Nature? The Resolutive-Compositive Method

Hobbes begins the *Leviathan* with three claims based on his earlier works that define the project for *Leviathan*. The first claim endorses a materialist view of the world: the world is but matter in motion.⁴ The second claim is that political institutions

1. The origins of the mutual advantage theory of justice are typically traced to Glaucon in Plato’s *Republic*, but the idea there is not developed in much detail. See Plato, *The Republic of Plato*, trans. Francis Macdonald Cornford (Oxford: Oxford University Press, 1945), 41–53.

2. Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*. Future references to *Leviathan* will use only the parenthetical reference style, indicating part, chapter, and section. Any of Hobbes’s other works will receive full citations.

3. See, for instance, Jean Hampton, *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1995); David P. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes* (Oxford: Clarendon Press, Oxford University Press, 1969); and Narveson, *The Libertarian Idea*.

4. Paragraph 1 of the Introduction. Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, 3.

are human creations.⁵ The third claim endorses a “resolutive-compositive” methodology: the way to understand something is to consider each of its parts individually and the manner in which they come together to form the object in question.⁶ From these three claims it follows that to understand the state, we should break it down into its component parts, *viz.* matter in motion, and consider how those parts affect one another when brought together to form the state. However, finding such a reduction to be practically impossible, Hobbes proposes to do the next best thing: to consider the nature of humans in a pre-state condition, and the manner in which they assemble into states.⁷ While a full reduction was practically impossible, Hobbes was a thorough reductionist, and this reductionism led him to posit a methodological continuity between the natural sciences and the studies of politics and ethics: the methods we use to understand natural phenomena are also the methods we should use to understand political phenomena.⁸

I will not investigate the first claim further as it would take us too far afield, but the second and third premise call for further attention if contractarianism is to have any relevance to normative ethics. While these three claims are explicitly focused on the state, they apply equally to ethics. Hobbes’s claim that the state is a human

5. Paragraph 2 of the Introduction. Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, 4.

6. Paragraph 2 of the Introduction. *ibid.* For more on Hobbes’s methodological claim, see Thomas Hobbes, “Philosophical Rudiments Concerning Government and Society (Translation of De Cive),” in *The English Works of Thomas Hobbes of Malmesbury*, ed. Sir William Molesworth, vol. 2 (London, UK: John Bohn, 1839), p. xiv; Douglas Jesseph, “Hobbes and the Method of Natural Science,” in *The Cambridge Companion to Hobbes*, ed. Tom Sorell (Cambridge: Cambridge University Press, 1996), 92–96; Hampton, *Hobbes and the Social Contract Tradition*, 7–8; John W.N. Watkins, *Hobbes’s System of Ideas* (London, UK: Hutchinson, 1965), 52–65; Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 1–4.

7. Hampton, *Hobbes and the Social Contract Tradition*, 12–13; Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 2; Hobbes, “Philosophical Rudiments Concerning Government and Society (Translation of De Cive),” p. xix–xx; Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, Introduction s. 3.

8. See *ibid.*, I.xv.40, I.ix.3.

creation can be recast as the claim that no one has authority over another human by nature—there are no natural slaves. If no one has authority to *command* another human, then the only remaining way to get another to behave as one would like is by *prayer* or by *counsel*. Command, prayer, and counsel will be examined in greater detail in section 2.1.5; for now I will note that to offer counsel to someone is to show that person that a certain course of action is in *that person's* interest, and that Hobbes contends that moral rules are counsels.⁹ While these moral rules describe what are effectively natural laws, providing knowledge of causal relations between actions and consequences, they derive their normativity from human desires. Thus moral rules are human creations in the sense that they rely on human interests for their force, and they bind us only because we take an interest in the good consequences that these bonds bring us. There are no moral requirements that apply to us independently of our desires. The reasoning for considering a state of nature will be revisited in section 3.1.3.

2.1.2 Human Nature

After explaining his method in “The Introduction,” Hobbes then provides a detailed and comprehensive account of the natural human condition in “Part I: Of Man,” exploring such topics as how humans sense things, how human passions are formed, what passions humans have, how humans use language, how humans reason, what science is, what religion is and why it arises. Only after covering such topics does Hobbes begin to consider how humans would get along in a stateless condition and what rational precepts would inform their interactions. While this thesis is not overly

9. For more, see section 2.1.5 on page 40. Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, (II.xxv.1–10); Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 68.

interested in all of these topics, some of the earlier elements offer crucial support for Hobbes's derivation of the moral laws. So while I will try to be more concise and selective than Hobbes, I will ask for your patience as I start at what can seem to be a rather early point in the story.

Different People Want Different Things at Different Times

Hobbes offers a materialist and mechanist explanation of humans, their senses, and their passions, which is heavily informed by his belief that all phenomena reduce to matter in motion. Humans have two kinds of movements: vital and voluntary (I.vi.2). Vital movements are the movements that our bodies do from (or before) birth, and continue to do until death, typically without conscious attention, for instance, the beating of our hearts, our circulation, and for the most part our respiration (I.vi.2). Our perceptions are caused by objects impacting our sense organs, thereby causing a motion that is conveyed to our brains and our hearts, where these motions either promote or inhibit our vital movements (I.i.4; I.vi.9–10). When our vital movements are inhibited, we experience pain or discomfort and desire to avoid the object that caused this inhibition; when our vital motions are promoted, we experience pleasure and seek out the object that caused the promotion (I.vi.9–10). Because each human body is slightly different, individuals differ in how they experience objects—either as pleasurable or painful (I.iv.24). Moreover, because each individual body changes as time passes, the same individual will find that some objects that once caused pain will later cause pleasure, and vice versa (I.iv.24; I.vi.6). Thus Hobbes thinks that humans will embrace diverse conceptions of the good, and that each individual's conception

of the good is likely to change over time (I.iv.24).¹⁰

Broadly construed, people want different things from what others want, and they will want different things over time. As Hobbes notes, this has implications for how we can proceed in reasoning. For as we experience things differently from each other, and differently from ourselves over time, we are bound to use evaluative terms like ‘good’ or ‘bad,’ differently.

[24] The names of such things as affect us, that is, which please and displease us, because all men be not alike affected with the same thing, nor the same man at all times, are in the common discourse of men of *inconstant* signification. . . . And therefore in reasoning a man must take heed of words which, besides the signification of what we imagine of their nature, have a signification also of the nature, disposition, and interest of the speaker, such as are the names of virtues and vices; for one man calleth *wisdom*, what another calleth *fear*; and one *cruelty*, what another *justice*; . . . And therefore such names can never be true grounds of any ratiocination. (I.iv.24)

[7] But whatsoever is the object of any man’s appetite or desire that is it which he for his part calleth *good*; and the object of his hate and aversion, *evil*; and of his contempt, *vile* and *inconsiderable*. For these words of good, evil, and contemptible are ever used with relation to the person that useth them, there being nothing simply and absolutely so, nor any common rule of good and evil to be taken from the nature of the objects themselves, but from the person of the man (where there is no commonwealth). . . .(I.vi.7)

10. By “conception of the good,” I mean only the things that people consider good, or that they desire. See section 2.1.2’s discussion of interest and desire.

So if we want to decide how we ought to act, or how we ought to structure our society, we ought not start by examining the common usage of evaluative terms like ‘justice.’

The Variability of Human Desires is Limited

Despite the claim that the objects of one’s desires will differ from the objects that others desire, and will differ from the former and future objects of one’s own desires, Hobbes still thinks that this variability has limits—that some things cannot be desired. This is because the system that Hobbes uses to explain the generation of desires is common to all humans. While we desire different things, our desires all come about by the same process: we perceive an object that either promotes or inhibits our vital motions, thereby causing us to desire the object if our vital motions are promoted, or to desire to avoid that object if it inhibits our vital motions. Given this story, it is impossible for one to desire an object that inhibits one’s vital motions, at least insofar as one is a healthy human being.¹¹ This gives Hobbes at least one human value that he can appeal to as universal: all humans must desire the preservation of their vital motions, and thus the preservation of their lives. Insofar as something can be shown to inhibit or promote an individual’s life, it has been shown to be in that individual’s interest. Here, “X is in the individual’s interest” is equivalent to “given the desires that the individual actually holds, the individual either actually desires X directly, or else the individual should desire X insofar as it is instrumental to one of the agent’s

11. Notions of health cause serious problems for Hobbes’s account, not only because they complicate the picture greatly, but because they cause problems for the derivation of moral principles, and threaten Hobbes’s ethical naturalism. For more on the threat to naturalism, see Jean Hampton, “Hobbes and Ethical Naturalism,” *Philosophical Perspectives* 6, *Ethics* (1992): 333–353. I will try to avoid appeals to health.

actual desires, whether or not the individual recognizes this connection.”¹² The individual can, thus, fail to desire what is in the individual’s interest, but this failure will reflect a failure of reason and an inconsistency in that individual’s desires. Given that individual’s own desires, that individual ought to desire objects that promote life, and to avoid objects that inhibit it.

People Always Want Something

The mechanist explanation of pain, pleasure, and desire also entails that individuals will perpetually desire new objects.

[58] *Continual success* in obtaining those things which a man from time to time desireth, . . . is that men call FELICITY; I mean the felicity of this life. For there is no such thing as perpetual tranquility of mind, while we live here; because life itself is but motion, and can never be without desire, nor without fear, no more than without sense. (I.vi.58)

Each time we perceive something, we will either want to pursue/maintain it or else to avoid it; we cannot be without desires until we stop perceiving, which is only likely to happen upon our deaths (I.viii.16). Hobbes adds to his justification for endless desires:

12. This definition of “in the individual’s interest” stands in contrast to the stronger definitions that many may have in mind when they use that phrase, many of which may posit an interest that does not depend in any way on the individual’s actual desires. Thus, many might want to say, for instance, that it is in an individual’s interest to experience fine culture whether or not the individual has any desires toward which experiencing fine culture would be instrumental. Hobbesians cannot make such claims. Indeed, I think that Hobbes cannot even say this about the desire for continued life: insofar as an individual does not desire to continue to live, and insofar as continuing to live is not instrumental to any other desire the individual holds, it cannot be shown that continuing to live is in that individual’s interest. All that Hobbes can say is that humans share a certain nature which makes it overwhelmingly probable that they will desire to continue to live. The rest of this thesis will talk about “desires” and “interests” as if they are interchangeable, and will mean things that are desired, or are instrumental to things desired by an individual.

Felicity is a continual progress of the desire, from one object to another, the attaining of the former being still but the way to the latter. The cause whereof is that the object of man's desire is not to enjoy once only, and for one instant of time, but to assure forever the way of his future desire. And therefore the voluntary actions and inclinations of all men tend, not only to the procuring, but also to the assuring of a continued life. . . . (I.xi.1)

While Hobbes's first argument claimed that we always want something more because we are always perceiving new things to which we cannot help but feel some desire or aversion, his second argument notes that we anticipate our future desires and try to guarantee in advance that they will be met. Thus, Hobbes thinks humans have "a perpetual and restless desire of power after power, that ceaseth only in death." (I.xi.2)

Practical Reason: Humans Seek to Maximize Pleasure and Minimize Pain

Anticipating Hume, Hobbes gives us a picture of human nature according to which human reason is "the slave of the passions."¹³ This slavery takes two forms. First, our abilities to reason well will be proportional to the usefulness such reasoning will have for us, which is proportional to our desire for power (I.viii.15–16). If we do not much care about preserving and improving our lives, we may not much care to become good at discerning helpful from harmful things.

Secondly, our passions direct our actions; reason can only serve to help fulfill the passions efficiently. For Hobbes, one's will is just the last desire one feels before acting (I.vi.49). One's will is shaped by deliberation—Hobbes's term for reasoning about

13. Hume, *A Treatise of Human Nature*, II.iii.3, p. 462. See also Hampton, *Hobbes and the Social Contract Tradition*, 19–20.

action—which seems to involve roughly four steps: (1) using reason or prudence, one predicts the consequences of different courses of action, (2) one imagines the consequences to determine the level of pleasure or pain that will accompany those consequences,¹⁴ (3) one weighs the anticipated pleasures and pains to find the balance for each course of action, and finally, though Hobbes does not explicitly state it, one can hardly help but add (4) one chooses the action expected to bring about the greatest amount of pleasure or the least amount of pain.

And because in deliberation the appetites and aversions are raised by foresight of the good and evil consequences and sequels of the action whereof we deliberate, the good or evil effect thereof dependeth on the foresight of a long chain of consequences, of which very seldom any man is able to see the end. But for so far as a man seeth, if the good in those consequences be greater than the evil, the whole chain is that which writers call *apparent* or *seeming* good. And contrarily, when the evil exceedeth the good, the whole is *apparent* or *seeming* evil; so that he who hath by experience or reason the greatest and surest prospect of consequences deliberates best himself, and is able, when he will, to give the best counsel unto others.

(I.vi.57)

Perhaps Hobbes does not include the prescription to choose that action which maximizes expected pleasure because he is attempting to describe the way that humans actually deliberate, and he acknowledges that humans can act irrationally (I.vi.53). Nevertheless, given all that Hobbes has said, it would seem that his conception of practical reason must include the prescription to act so as to bring about the best

14. Hobbes calls the pleasure arising from anticipated consequences “*pleasures of the mind*” (I.vi.12).

anticipated consequences; otherwise, it is unclear what it is about human deliberation and action that could make the action irrational.¹⁵

Human Equality

The final key piece of Hobbes's account of human nature is human equality. This consideration will lead us very naturally to considering what life is like for humans in a state of nature, and what principles would govern their interactions in such a state. Moreover, we have already begun talking about human equality while talking about human reason. Not only does Hobbes think that humans are roughly equal in prudence (I.xiii.2), he also thinks that they are roughly equal in physical ability in a rather peculiar way.

[1] Nature hath made men so equal in the faculties of body and mind as that, though there be found one man sometimes manifestly stronger in body or of quicker mind than another, yet when all is reckoned together the difference between man and man is not so considerable as that one man can thereupon claim to himself any benefit to which another may not pretend as well as he. For as to the strength of body, the weakest has strength enough to kill the strongest, either by secret machination, or by confederacy with others that are in the same danger with himself.

(I.xiii.1)

This may sound like a very strange notion of equality, and to some extent it is, yet it can sound more plausible when we connect it with Hobbes's political naturalist

15. Jean Hampton and John Elster also think it reasonable to attribute to Hobbes a traditional account of practical reason according to which one ought to act to maximize expected happiness and minimize expected pain. See Hampton's discussion of Elster in Hampton, *Hobbes and the Social Contract Tradition*, 16.

assumption discussed in subsection 2.1.1.¹⁶ As Jean Hampton notes, Hobbes is discussing human equality in connection with political claims of entitlement.

Hobbes's assumption of human equality may strike modern readers as implausibly strong. But it is important to see that Hobbes does not need to say or even intend to say that there are no large differences in the intellectual and physical abilities among human beings; rather, he is trying to make the point that our differences are never so great as to make some of us natural slaves and others natural masters in a state of nature. Whatever differences in ability or strength or talent differentiate us, they have no *political* significance. . . . [Hobbes's] point is that because no one is so superior mentally or physically as to be able to establish rulership quickly and securely over everyone else *without* their cooperation, we must be treated from a political standpoint as equal to one another. . . .¹⁷

While Hobbes's claim here is explicitly political, and is interpreted in a political way by Hampton, it holds equally of our moral relations. Because of our rough equality, no one is naturally entitled to have his or her interests recognized and respected by others. If we are to justify moral duties, we cannot do so by appeal to any natural status.

To conclude this section, humans have different desires from each other, and each individual's desires will change over time. However, given the way our desires are formed, we can never want anything detrimental to our life. Moreover, we will always want things; we can never be satisfied. Our desires lead us to want to reason well, as good reasoning abilities will give us power insofar as we can use causal knowledge

16. Page 19.

17. Hampton, *Hobbes and the Social Contract Tradition*, 25–26.

to shape the world for our gain. And desires drive our practical reasoning: we seek to best fulfill our desires by our actions. When it comes to our ability to predict the effects of causes without linguistic reasoning, we are all about equally good at it; and when it comes to killing each other, we are all about equally capable at that too.¹⁸ This concludes Hobbes's account of the constituent matter and the authors of the state; the next stage in the process is to examine how and why they interact to form the state.

2.1.3 The Problem of the State of Nature: War

After presenting his account of human nature, Hobbes sets out to explain how humans would interact in a state of nature—a state devoid of any political authority—given their nature. This marks one primary division in Hobbes interpretation: systematic interpreters, whom I follow here, believe that Hobbes derives his moral principles from his account of human nature without appealing to any natural moral facts, whereas non-systematic interpreters believe that Hobbes uses natural moral concepts that do not come from his account of human nature.¹⁹ Given that my primary aim is to identify the compelling elements of social contract theory, and that the non-systematic interpretation produces an unconvincing interpretation, I will dismiss the

18. Note that there is nothing particularly antisocial or barbaric about human nature on Hobbes's account. As the next section will show, the problems of the state of nature arise from human nature *in a particular situation*. Thus, critics like La Barbera who claim that Hobbes's account of human nature is horrifying and sociopathic are mistaken. (Christopher La Barbera, *States of Nature: Animality and the Polis* (New York: Peter Lang, 2012), 41–44).

19. For non-systematic interpretations, see: A. E. Taylor, "The Ethical Doctrine of Hobbes," in *Hobbes Studies*, ed. Keith Brown (Oxford: Blackwell, 1965), 35–55; Howard Warrender, *The Political Philosophy of Hobbes* (Oxford: Clarendon Press, 1957); Bernard Gert, *Man and Citizen* (Atlantic Highlands, N.J.: Humanities Press, 1978). For systematic interpretations, see: Hampton, *Hobbes and the Social Contract Tradition*; Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*; R. S. Peters, *Hobbes* (Harmondsworth: Penguin, 1956); Thomas Nagel, "Hobbes's Concept of Obligation," *Philosophical Review* 68 (1959): 68–83; Watkins, *Hobbes's System of Ideas*.

non-systematic interpretation without further attention.

How will humans interact in the state of nature? We recall that each individual wants things, and always will want things. Given that human desires are potentially limitless, and that the materials in the world are limited, it is bound to arise that multiple individuals want one and the same thing. This problem is aggravated by the fact that Hobbes thinks that some humans also desire relative status, for instance, wanting to be richer than anyone else. And when two individuals want the same object, given human equality, neither of them has a natural claim to that thing (or anything for that matter). Here we have the first cause of war.

[3] Competition of riches, honour, command, or other power, inclineth to contention, enmity, and war; because the way of one competitor to the attaining of his desire is to kill, subdue, supplant, or repel the other.
(I.xi.3).

Not only does human equality entail that neither party has a natural entitlement or claim that the other should yield, so that there is no natural solution to the problem, but equality also actively contributes to the problem by encouraging each party to fight.

[3] From this equality of ability ariseth equality of hope in the attaining of our ends. And therefore, if any two men desire the same thing, which nevertheless they cannot both enjoy, they become enemies, and in the way to their end, which is principally their own conservation, and sometimes their delectation only, endeavor to destroy or subdue one another. And from hence it comes to pass that, where an invader hath no more to fear than another man's single power, if one plant, sow, build, or possess a

convenient seat, others may probably be expected to come prepared with forces united to dispossess and deprive him, not only of the fruits of his labour, but also of his life or liberty. And the invader again is in the like danger of another. (I.xiii.3)

Thus, in a state of nature devoid of political institutions like a system of legally enforceable property rights, human rapaciousness and human equality combine to bring it about that more than one person desires the same thing, and that these people have no recourse but to fight over the thing in question. A keen and skeptical reader will surely ask: “Could they not just agree on some mutually advantageous principles?” That is the beginning to Hobbes’s solution concept for war, and we will come to that, but for the moment Hobbes is examining the problem before examining the solution.

Hobbes’s discussion of competition leads to his discussion of diffidence, or distrust, as a second cause of war. Anyone who possesses a good that others might find valuable has reason to anticipate, given competition, that others will attempt to take that good by force, potentially killing the current possessor.

[4] And from this diffidence of one another, there is no way for any man to secure himself so reasonable as anticipation, that is, by force or wiles to master the persons of all men he can, so long till he sees no power great enough to endanger him. And this is no more than his own conservation requireth, and is generally allowed. (I.xiii.4)

Because humans not only want things now, but also want to secure and provide for their future desires, they will find it in their interest to preemptively attack anyone who might attack them in the future.

Hobbes adds a third cause of war: glory. After explaining the genesis of human desires, Hobbes catalogues the human passions, among which is glory. Glory, in itself, can be healthy, and consists in the feeling of pleasure one gets from remembering great deeds one has done (I.vi.39). Yet in some individuals, the delight of glory can be excessive.

Also, because there be some that taking pleasure in contemplating their own power in the acts of conquest, which they pursue farther than their security requires, if others (that otherwise would be glad to be at ease within modest bounds) should not by invasion increase their power, they would not be able, long time, by standing only on their defence, to subsist. And by consequences, such augmentation of dominion over men being necessary to a man's conservation, it ought to be allowed him. (I.xiii.4)

Some people derive excessive pleasure from reflecting upon their triumphs in battle, and this leads them to attack others for that reason alone.²⁰ Not only do such people contribute directly to the war of all against all, but they further reinforce diffidence as a cause of war: now, one must be suspicious of others not only because they might want to steal one's possessions, but also because they might want to attack one for the sheer glory of victory.

Thus, not only does Hobbes think that the state of nature will be a state of war of each against each other, but that this war will be accompanied by numerous indirect

20. How much room is there, in Hobbes's account of human nature, for Hobbes to criticize someone for taking excessive delight in glory? Practical reason is a slave of the passions; how could a passion be criticized as irrational? This question will be further examined in section 3.1.2 on page 82. See also Susan Dimock's "Two Virtues of Contractarianism," where she explains that instrumental accounts of rationality like Hobbes's can call for a revision of our goals and desires, for example, calling on us to abandon goals and desires that would be impossible or very costly to fulfill. (Susan Dimock, "Two Virtues of Contractarianism," *Journal of Value Inquiry* 37 (2003): 199–400.)

“incommodities.”

[9] Whatsoever therefore is consequent to a time of war, where every man is enemy to every man, the same is consequent to the time wherein men live without other security than what their own strength and their own invention shall furnish them withal. In such condition there is no place for industry, because the fruit thereof is uncertain, and consequently, no culture of the earth, no navigation, nor use of the commodities that may be imported by sea, no commodious building, no instruments of moving and removing such things as require much force, no knowledge of the face of the earth, no account of time, no arts, no letters, no society, and which is worst of all, continual fear and danger of violent death, and the life of man, solitary, poor, nasty, brutish, and short. (I.xiii.9)

Not only will individuals be at constant risk of attack, but because of this risk, it will be irrational for anyone to attempt any task that takes much effort and that is not immediately conducive to defence, preemptive attack, and survival. The probability that one will get to enjoy the benefits of any effort are slim; thus, in order for an activity to be, on balance, an anticipated good, the costs of the activity will have to be very low, and the benefits, however unlikely, will have to be quite considerable.

2.1.4 The Right of Nature

This state of war leads individuals to claim a natural right to all things. Hobbes has a rather idiosyncratic notion of “right” which is not immediately obvious from his definition of the term.

[1] the RIGHT OF NATURE, which writers commonly call *jus naturale*, is

the liberty each man hath to use his own power, as he will himself, for the preservation of his own nature, that is to say, of his own life, and consequently of doing anything which, in his own judgment and reason, he shall conceive to be the aptest means thereunto. (I.xiv.1)

Gauthier offers a formal and a material interpretation of right: the formal definition is that one has a right to do whatever right reason says one should do, and the material definition, which builds in Hobbes's account of human desires, is that one has a right to do whatever preserves one's life.²¹ This distinction may help to clarify the confusion in the quotation: Hobbes endorses the formal interpretation of right, and because of his account of human nature and rationality, Hobbes thinks that the formal interpretation entails the material interpretation. Nevertheless, the peculiar feature of Hobbes's account of right is not yet apparent; it becomes obvious when Hobbes discusses the right to all things.

[4] And because the condition of man (as hath been declared in the precedent chapter) is a condition of war of everyone against everyone (in which case everyone is governed by his own reason and there is nothing he can make use of that may not be a help unto him in preserving his life against his enemies), it followeth that in such a condition every man has a right to everything, even to one another's body. And therefore, as long as this natural right of every man to everything endureth, there can be no security to any man. . . . (I.xiv.4)

Here the peculiar aspect of Hobbes's conception of right comes to the fore: for Hobbes, to claim that X has a right to do Y is not to claim that anyone has an obligation not

21. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 31–33, 49–52.

to obstruct X's doing of Y. Natural rights have no correlative duty. To say "X has a natural right to Y" is but to say that X has no obligation not to Y. Or at least, this is half of Hobbes's conception of right.

Other scholars follow Gauthier in stressing the connection between right and reason for Hobbes, so that "X has a natural right to Y" entails that X has no obligation not to Y *and* Y is not counter to right reason for X—that is, it is not irrational for X to Y.²² This allows for some benefits, like being able to claim that the laws of nature do not conflict with the rights of nature,²³ but I find it problematic for a couple of reasons.

First, this interpretation seems inconsistent with what Hobbes says in other places, as Gauthier recognizes:²⁴

For though they that speak of this subject use to confound *jus* and *lex* (*right* and *law*), yet they ought to be distinguished, because RIGHT consists in liberty to do or to forbear, whereas LAW determineth and bindeth to one of them; so that law and right differ as much as obligation and liberty, which in one and the same matter are inconsistent. (I.xiv.3)

Hobbes claims that having a right gives one the ability to choose between doing or not doing something. In order to make sense of this, those like Gauthier, who connect right with reason, claim that we have a right when none of the options under consideration are forbidden by right reason. Say Joe is considering what to do at the moment, and he can either attack Sam or not attack Sam. If both of those options are concordant with right reason—if both of them are anticipated goods—then Joe

22. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 30–35; Hampton, *Hobbes and the Social Contract Tradition*, 35.

23. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 39.

24. *ibid.*, 33–34.

has a right to attack Sam and to not attack Sam. But if attacking Sam is not an anticipated good—if attacking Sam is expected to go very badly because Sam is a well-armed giant of a man—then Joe has a right to not attack Sam, but has no right to attack Sam. But here Joe’s right does not consist in the “liberty to do or forbear,” and looks rather more like an obligation to not attack Sam.²⁵ So this interpretation seems inconsistent with what Hobbes says, but, given that Hobbes contradicts himself (or seems to) in a number of places, it is important to press further.

Gauthier’s interpretation is also problematic because it obviates Hobbes’s conception of natural law. I want to save natural law for section 2.1.5, but for the moment we can note that natural laws, for Hobbes, command that which is required by reason. According to Gauthier’s interpretation, for any action it will be the case that it is forbidden by law if and only if one does not have a right to perform the act, and both for the exact same reason: because the action is a seeming bad and thus proscribed by right reason. Thus, if one follows Gauthier’s interpretation of right and allows right to play a prescriptive role based on its connection with reason, then one will have to admit that natural laws add nothing—they simply restate the prescriptions given by rights.

Most importantly, this interpretation of right is problematic because it would make Hobbes’s claim that individuals have a natural right to all things false. Like the last objection, this objection requires an anachronistic discussion of that which will come later. Hobbes wants to partially solve the problems of the state of nature by showing that it is rational to constrain one’s behavior. Yet if it is rational to

25. Of course, it does not look like an obligation not to attack Sam given Hobbes’s definition of the term “obligation,” but when Hobbes says that right and law “differ as much as obligation and liberty,” he seems to be using “obligation” and “liberty” in ways that do not fit with his definition of the terms (I.xiv.3). I suspect that in this sentence he is using “obligation” and “liberty” in accordance with general usage, and in (a harmless) violation of his definitions of the terms.

constrain one's behavior, then individuals do not naturally have an unlimited right to all things. If the right of nature is tied too closely to reason, then the problems of the state of nature do not arise insofar as those problems arise from irrational acts.²⁶ All of this leads us to another interpretive difficulty: some scholars emphasize the role of the right to all things as a cause of war, while others claim that it is the beginning of the solution to the state of war. Hobbes himself seems to be conflicted on this issue, presenting the equal right to all things as a cause of war in *De Cive*, and as either the result of war or as the beginning of a solution to war in *Leviathan*.²⁷

For the sake of moving forward, I will opt for clarity. I think that severing the concept of right from reason will, at the very least, afford us a measure more clarity by allowing us to talk about the two concepts separately. Thus, I think that Hobbes's claim that individuals have an unlimited natural right to all things is equivalent to his earlier claim that in "this war of every man against every man, this also is consequent: that nothing can be unjust. The notions of right and wrong, justice and injustice, have there no place." (I.xiii.13). I think Hobbes is embracing the account of right I presented earlier: if X has a right to Y, then it is not unjust for X to Y. Given Hobbes's account of human nature, if one has an unlimited right to all things, then one will do whatever one finds rational, which means that one will do whatever is life-preserving and pleasure-giving. Adding to this Hobbes's claim that in the state of nature there is nothing that one would not find conducive to one's ends, Hobbes's claim that individuals would feel rationally compelled to attempt to gain control over all things follows. Thus this narrower definition of right still allows Hobbes to derive all of the implications that he could derive from the reason-connected definition of right;

26. The connection between reason, the laws of nature, and the ills of the state of nature is drawn out further in section 2.1.5 on page 40.

27. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 34–35.

it simply forces him to keep right and reason distinct for the sake of clarity. I think this move is akin to Gauthier's attempt to separate material from formal concepts, but I think that Gauthier left a material assumption in his formal definition of right, which I have now removed. This revision of the concept of 'right' does not allow for the claim that laws do not limit rights, but that seems to be a claim that Gauthier was more wedded to than was Hobbes.²⁸

The upshot of this discussion is that Hobbes thinks that in the state of nature each individual has an unlimited right to all things, meaning that there is nothing such a person could do that would be unjust. No one is entitled to any specific treatment by anyone else. Not only does each person have a right to all things, but there is nothing that each person would not find useful in the defence of her or his life and happiness. So not only does each person have a right to all things, but each person also has an interest in taking all things, because the more things an individual has, the more power that individual has. And power is needed to preserve one's life and one's holdings against others. Moreover, what matters is *relative* power: Joe's power is good for nothing if there is a more powerful group around, because that more powerful group will find what Joe has to be valuable and take it. Thus each individual will constantly be striving to be *most* powerful.

28. While I think this interpretation of "right" affords the greatest clarity consistent with what Hobbes says, I must admit that it is not always consistent with what Hobbes says. For instance, Hobbes claims that "he which performeth first does but betray himself to his enemy, contrary to the right (he can never abandon) of defending his life and means of living" (I.xiv.18). Here Hobbes clearly implies that 'right' is more than just permission; it has some prescriptive force. While my interpretation is not consistent with such claims, I still prefer my interpretation (a) because such claims can be rephrased in terms of rationality instead of right with no loss, so that Hobbes will still get the conclusions he wants once the claims are rephrased, and (b) because the interpretation of right connected with reason, on the other hand, does interfere with Hobbes's conclusions, like his conclusion that there is a universal equal right to all things in the state of nature, as discussed above.

2.1.5 The Laws of Nature

After introducing the unlimited natural right to all things, Hobbes presents the laws of nature. Once again, Hobbes has a rather idiosyncratic notion of “law.”

A Law of Nature (*lex naturalis*) is a precept or general rule, found out by reason, by which a man is forbidden to do that which is destructive of his life or taketh away the means of preserving the same, and to omit that by which he thinketh it may be best preserved. (I.xiv.3)

Hobbes admits that “law” is, indeed, a misnomer: what he is talking about are the conclusions or theorems of practical reason. Hobbes distinguishes three kinds of directives, or statements recommending a certain action: command, prayer, and counsel (II.xxv.1–10). When Joe commands or prays that Sam do Y, Y is an expected good for Joe, but may not be for Sam. Command is differentiated from prayer by authority: Joe can only command Sam if he has authority over Sam, otherwise Joe can only pray and hope that Sam does Y out of kindness or some other sympathetic passion. On the contrary, when Joe counsels Sam to do Y, Y is in Sam’s interest, and Joe is explaining to Sam how Y is in Sam’s interest. Natural laws, for Hobbes, are counsels.²⁹ There is no one in the state of nature with the authority to command another, and prayer is an unreliable way to get someone to do something, so the only way to get anyone to do a particular action is by counsel—by showing it to be in the interest of the actor.

What differentiates the laws of nature from other counsels is that the laws have a measure of objectivity and universality to them that other forms of counsel lack. Joe

²⁹ Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 68; Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, II.xxv.1–10.

may be able to show Sam that planting crops is in Sam's interest, but Joe might not succeed when he tries that same line of reasoning on Samantha. This follows from the variability of human desires: Sam may have some desire that makes planting crops an anticipated good for Sam, yet Samantha may lack that desire. Not only do the laws of nature appeal only to those desires that are shared by all humans, they also reconcile each individual's desires with those of others. A counsel could appeal to a universal desire and say something like, "insofar as you value your life, you should kill everyone else." The laws do more than this; they provide rules for satisfying everyone's desires in a manner that is desirable to each individual.

At no stage of his argument does Hobbes lose sight of the viewpoint of the individual. Although his moral system provides a common viewpoint for all men, yet each man must see what is common as the outcome of his own particular concern with his own preservation and well-being. Thus each new step in the argument is justified to each man in terms of his own good.... The laws of nature are maxims for the preservation of the individual, but they hold for every individual. If each man preserves himself, then all men—society—is preserved. But the basic consideration is always the preservation of the individual.³⁰

Thus, the counsel "if you value your life, kill everyone" may appeal to each individual's desires, but it is mutually disadvantageous: insofar as each individual adopts this strategy, their actions are mutually self-defeating. The laws of nature provide mutual advantage. They promise to improve the life of each individual, but they do so by improving everyone else's life as well.

30. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 90.

So while these are natural laws, they are not some kind of strange metaphysical entity or irreducibly moral fact lacking justification. In keeping with his scientific method, Hobbes considers the laws of nature to be true causal scientific knowledge: the laws tell us which consequences follow from which actions. They speak to us insofar as the consequences they discuss are of the utmost importance to each of us.³¹ In one sense, these laws are not artificial or conventional: they capture what Cartwright calls a “capacity”—a steady contribution or connection between certain human actions and certain outcomes.³² In the same way that heating water to one-hundred degrees Celsius at sea level will cause the water to boil, following the laws of nature will cause peace, and breaking them will cause war. Human reason is needed to identify these laws, and human reason employs the laws to determine the most effective means to what we desire.³³

Yet the first two laws of nature direct us to take on further conventional duties through obligation. The first law directs each individual to seek peace:

[I]t is a precept, or general rule, of reason *that every man ought to endeavor*

31. To put the point in Kantian terms, the laws of nature are “assertorial” hypothetical imperatives: “There is *one* end, however, which may be assumed to be actually such to all rational beings. . . and therefore, one purpose which they not merely *may* have, but which we may with certainty assume that they all actually *have* by a natural necessity, and this is *happiness*. The hypothetical imperative which expresses the practical necessity of an action as means to the advancement of happiness is Assertorial. We are not to present it as necessary for an uncertain and merely possible purpose, but for a purpose which we may presuppose with certainty and *à priori* in every man, because it belongs to his being.” (Immanuel Kant, *Fundamental Principles of the Metaphysics of Morals*, trans. T. K. Abbott (Amherst, NY: Prometheus Books, 1988), 43) Kant also calls these imperatives “*counsels* of prudence. . . Counsels, indeed, involve necessity, but one which can only hold under a contingent subjective condition, viz. they depend on whether this or that man reckons this or that as part of his happiness. . .” (ibid., 44).

32. See Nancy Cartwright, *Nature’s Capacities and their Measurement* (Oxford: Clarendon Press, Oxford University Press, 1994).

33. Thus, there is no contradiction in Hobbes’s claim that the laws are natural and rules of reason, *pace* La Barbera. (La Barbera, *States of Nature: Animality and the Polis*, 46–47). For more on the naturalism of Hobbesian contract theory, and its ability to avoid the naturalistic fallacy and Moore’s “open question” objection, see Dimock, “Two Virtues of Contractarianism.”

peace, as far as he has hope of obtaining it, and when he cannot obtain it, that he may seek and use all helps and advantages of war. (I.xiv.4)

And the second law directs us to lay down our rights in pursuit of peace.

[5] From this fundamental law of nature, by which men are commanded to endeavor peace, is derived the second law: *that a man be willing, when others are so too, as far-forth as for peace and defence of himself he shall think it necessary, to lay down his right to all things, and be contented with so much liberty against other men as he would allow other men against himself. (I.xiv.5)*

From there, Hobbes goes on to catalogue seventeen more laws of nature: perform contracts made (I.xv.1); do not make gift-givers regret their gift (I.xv.16); be accommodating (I.xv.17); forgive offenders who seek forgiveness (I.xv.18); punish only to the extent that punishment is helpful to deterrence (I.xv.19); do not publicly disrespect others (I.xv.20); recognize all humans as equal (I.xv.21); do not hold out for more than you need (I.xv.22); deal equally when arbitrating a dispute (I.xv.23–4); share indivisible objects in common (I.xv.25); assign indivisible objects by natural lot (I.xv.26); natural lot belongs either to the first born or the first possessor (I.xv.28); do not attack mediators (I.xv.29); use arbitration to settle disputes (I.xv.30); a disputant cannot arbitrate in his or her own case (I.xv.31); arbitrators must not stand to gain from their decisions (I.xv.32); and finally, arbitrators ought to give equal credence to disputants and use witnesses to settle disagreements (I.xv.33).

The particular laws are not of particular relevance here beyond the third law. The first three laws lay the foundations for further conventional obligations: one should cooperate with others, but only insofar as they are willing to do the same. If you lay

down your rights when others are refusing to do the same, your action will contribute neither to peace nor to ameliorating your life; in fact, you will not live long if you make such a unilateral renunciation of right. But if others are laying down their rights as well, then your renunciation will cause the desired effects. Thus the first three laws call for further conventional laws that derive their force in part from the convention and in part from the natural laws.

2.1.6 Contract and Cooperation in the State of Nature

There are two ways to lay down one's right: by transferring it or by renouncing it (I.xiv.6–7). Neither renouncing nor transferring one's right gives anyone else any new right that they did not have before, since everyone already has an unlimited right to all things (I.xiv.6). When one renounces one's right to possess an object, for instance, one commits oneself to not interfering if anyone decides to exercise their right to all things and take possession of that object (I.xiv.7). When one transfers one's right to an object, one commits to not interfere with a specific person's attempts to possess that object (I.xiv.7). One lays down one's right by contract, and the contract obliges one (I.xiv.7). Employing Gauthier's formal-material distinction can help us gain a bit more clarity here: formally, to contract is to indicate by words or gestures that one lays down one's right, but once one takes Hobbes's conception of human nature into account, to contract, materially, is to transfer one's right to another person in exchange for some right against that person. This is simply because no one would lay down a right unless they stood to gain from doing so, and that will only happen if there is a mutual transfer of right (I.xiv.9).

We can now consider the challenge that keen readers brought up earlier: it seems

like the state of nature need not be a state of war insofar as we can make a mutually advantageous contract with another when we discover that we want the same objects. Hobbes jumped from the fact that people will want the same objects to the conclusion that they will fight over them, but it seems more plausible that they will try to settle on mutually advantageous agreements.

There is perhaps no other issue where Hobbes seems to contradict himself so severely and repeatedly as the issue of contracting in the state of nature. This seeming contradiction arises because Hobbes is being pulled in opposite directions: on the one hand, everything he has said about human nature and the laws of nature—especially the law requiring that contracts be upheld—indicates that humans should be capable of forming and upholding contracts in the state of nature; but on the other hand, Hobbes wants the state of nature to be a state dominated by violence in order for his argument for the creation of an absolute sovereign to work. Moreover, although his argument for the institution of the sovereign requires that the state of nature be sufficiently dysfunctional, it also requires humans to be capable of forming and fulfilling contracts to a certain extent. So even Hobbes’s political aims pull him in both directions with respect to the ability to contract in the state of nature.³⁴

Thus we find passages where Hobbes claims that the keeping of contracts is impossible in the state of nature.

For the laws of nature. . . of themselves, without the terror of some power to

34. It is perhaps this tension that led La Barbera to his interpretation of Hobbes’s argument as positing that human nature is brutal and that nature and animality must be suppressed. See Chapter III in La Barbera, *States of Nature: Animality and the Polis*, 37–55. But the kind of conversion from a ‘natural human nature’ to a ‘civilized human nature,’ and the antagonism between the two natures, that La Barbera reads into Hobbes, renders an implausible interpretation of Hobbes’s arguments that Hampton discusses and rejects. (Hampton, *Hobbes and the Social Contract Tradition*, 208–220) The rest of this section avoids this conversion interpretation.

cause them to be observed, are contrary to our natural passions, that carry us to partiality, pride, revenge, and the like. And covenants without the sword are but words, and of no strength to secure a man at all. Therefore notwithstanding the laws of nature (which every one hath then kept, when he has the will to keep them, when he can do it safely), if there be no power erected, or not great enough for our security, every man will, and may lawfully rely on his own strength and art, for caution against all other men. (II.xvii.2)

And yet, we also find passages where Hobbes implies that contracts can be upheld in the state of nature.

And when a man hath in either manner abandoned or granted away his right, then is he said to be OBLIGED or BOUND not to hinder those to whom such right is granted or abandoned from the benefit of it; and [it is said] that he *ought*, and that it is his DUTY, not to make void that voluntary act of his own, and that such hindrance is INJUSTICE, and INJURY, as being *sine jure* [without right], the right being before renounced or transferred. (I.xiv.7).

Moreover, Hobbes insists that coerced contracts are valid in the state of nature.

Covenants entered into by fear, in the condition of mere nature, are obligatory. For example, if I covenant to pay a ransom, or service, for my life, to an enemy, I am bound by it. For it is a contract wherein one receiveth the benefit of life; the other is to receive money, or service, for it; and consequently, where no other law (as in the condition of mere nature) forbiddeth the performance, the covenant is valid. (I.xiv.27)

Hobbes then tries to qualify the extent to which contracts are valid in the state of nature. Thus, he distinguishes covenants from contracts, where covenants are a subset of contracts in which one party performs her half of the deal before the other party needs to decide whether or not to perform his half. Hobbes thinks that when one is the second party to perform in a covenant, if the other party has performed then the contract is valid and thus one ought to perform. But Hobbes does not think that anyone is ever required to perform first, and in fact thinks that performing first in the state of nature is unwise.

For he that performeth first has no assurance the other will perform after, because the bonds of words are too weak to bridle men's ambition, avarice, anger, and other passions, without the fear of some coercive power; which in the condition of mere nature, where all men are equal and judges of the justness of their own fears, cannot possibly be supposed. And therefore, he which performeth first does but betray himself to his enemy, contrary to the right (he can never abandon) of defending his life and means of living. (I.xiv.18)

So while it is not (ever) rational to perform one's part first, once one party has performed, the other party is obligated to perform. This is where we encounter "the Fool."

Hobbes considers an objection to his claim that a covenant is binding on the second party to perform once the first part has done so. Charitably, Hobbes calls the objector who presents this argument "the Fool."

The fool hath said in his heart 'there is no such thing as justice... [and] every man's conservation and contentment being committed to his own

care, there could be no reason why every man might not do what he thought conduced thereunto, and therefore also to make or not make, keep or not keep, covenants was not against reason, when it conduced to one's benefit.' He does not therein deny that there be covenants, and that they are sometimes broken, sometimes kept, and that such breach of them may be called injustice, and the observance of them justice; but he questioneth whether injustice, taking away the fear of God. . . , may not sometimes stand with that reason which dictateth to every man his own good; and particularly then, when it conduceth to such a benefit as shall put a man in a condition to neglect, not only the dispraise and revilings, but also the power of other men. . . . '[I]f it be not against reason, it is not against justice; or else justice is not to be approved for good.' (I.xv.4)

The Fool claims that it is sometimes in one's interest, when one is the second party to perform in a covenant, to violate the agreement. And since, for Hobbes, the only thing that can motivate someone to act in a given way is interest, the laws of nature cannot proscribe the violation of contracts without sacrificing their motivating force in some cases. That is, if we say that the laws forbid the breaking of contracts in all cases, and yet we find ourselves in a case where it is in our interest to violate a contract, then we have no reason to obey or recognize the law of nature. Otherwise, we must say that the laws of nature permit the violation of contracts. What we decide to say does not matter much to the Fool; either way, he will do what is in his interest by breaking the contract.

Hobbes's response is particularly interesting to his modern interpreters because what he demonstrates differs markedly from what he thinks he has demonstrated

(and likely all he intended to demonstrate). Firstly, Hobbes grants the Fool's point in part: as Hobbes's notion of 'right' was peculiar, so is his concept of "obligation." This peculiarity arises, again, because the only considerations that can motivate an individual to act are considerations of that individual's pain, pleasure, and continued life. Thus, for Hobbes, Joe has an obligation to do something for Sam if and only if it is the case both that Joe contractually transferred a right of his to Sam, and that fulfilling the agreement is in Joe's interest.³⁵ This leads Gauthier to criticize Hobbes, claiming that Hobbes does not actually have any relevant moral conception of obligation at all.³⁶ Nevertheless, this appears to be Hobbes's understanding of "obligation," which is confirmed in the first part of his response to the Fool: Hobbes concedes that one is under no obligation to uphold one's end of the bargain *unless* either there is a sovereign power to enforce contracts, or the other party has already performed.

For the question is not of promises mutual where there is no security of performance on either side (as when there is no civil power erected over the parties promising), for such promises are no covenants, but either where one of the parties has performed already, or where there is a power to make him perform, there is the question whether it be against reason, that is, against the benefit of the other to perform or not. And I say it is not against reason. (I.xv.5)

Thus, whenever one's obligation would run counter to one's interest, one is not under obligation, and the contract is not valid.

35. Hampton, *Hobbes and the Social Contract Tradition*, 55–56.

36. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 60–61.

Nevertheless, Hobbes identifies two cases where contracts are valid: where *either* there is a coercive state to punish the violation of contract, *or* where the other party has already performed.³⁷ Setting aside the former case for now, Hobbes argues that it is rational to perform your part of a contract if the other party has already performed:

[I]n a condition of war wherein every man to every man (for want of a common power to keep them all in awe) is an enemy, there is no man can hope by his own strength or wit to defend himself from destruction without the help of confederates (where everyone expects the same defence by the confederation that anyone else does); and therefore, he which declares he thinks it reason to deceive those that help him can in reason expect no other means of safety than what can be had from his own single power. He, therefore, that breaketh his covenant, and consequently declareth that he thinks he may with reason do so, cannot be received into any society that unite themselves for peace and defence but by error of them that receive him; nor when he is received, be retained in it without seeing the danger of their error;. . . and therefore, if he be left or cast out of society, he perisheth. . . . (I.xv.5)

Some think that Hobbes is cheating here in his response to the Fool. The Fool says: if I am in a situation where it is in my interest to cheat then I should cheat. And

37. Hampton emphasizes the disjunction that appears in I.xv.5, quoted above. In the Latin version, which many argue antecedes the English, this passage is conjunctive: contracts are valid where there is a sovereign *and* the other party has performed. We need not rely on the claim that the English version came later to argue that Hobbes actually believed the disjunctive English version and not the conjunctive Latin version, for in both versions Hobbes goes on in the same section to argue explicitly that contracts are valid in the state of nature. Thus, the Latin version is inconsistent, and appears to be a mistake. Hampton, *Hobbes and the Social Contract Tradition*, 64–65. For the Latin version, see Curley’s footnote 5 to I.xv.5, Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, 91.

Hobbes's response is that it is never in one's interest to cheat. In the state of nature, no one has sufficient strength to maintain his or her life by force alone; one must seek out confederates and join together in defence pacts.³⁸ One can imagine the Fool responding: yeah, but if it was in my interest, I should cheat. There is a sense in which Hobbes and the Fool are talking past one another.³⁹ Hobbes must agree with the Fool: obligation against interest is no obligation at all. However, the Fool ought to recognize that this concession does not imply that one should go around doing as one pleases, for, if Hobbes is right, the circumstances in which it is in one's interest to cheat will (almost) never arise.

For now, let us grant that Hobbes has shown that it is rational to uphold one's end of a contract whenever the other party has already performed. Hobbes thinks that this is the end of the discussion; however, modern interpreters have argued that, in fact, Hobbes's argument shows that it is rational to form contracts, to perform first, and to perform second, which is much more than Hobbes thought he had demonstrated, and much more than he wished to demonstrate. Consider the position of the party who is considering performing first, and who reasons as follows: "If I perform first, then it will be rational for the other party to perform, because it is (almost) always rational for someone to do their part when they are the second to perform. I prefer

38. This kind of reasoning is remarkably similar to Nozick's, and leads Hampton to claim that, despite the fact that he does not make explicit appeal to a contract, Nozick is nonetheless a contractarian. Hampton, *Hobbes and the Social Contract Tradition*, 168–169; Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974).

39. Hobbes's response is a case of the fallacy of denying the antecedent. But I do not think that this gets us to the heart of the disagreement between Hobbes and the Fool; I think that they disagree primarily over the number of cases where the antecedent will be true—the number of cases where it will be in the Fool's interest to cheat. I contend that Hobbes should accept that the Fool should cheat when it is in the Fool's interest to do so; Hobbes's argument should be that it will never—or at least almost never—be in the Fool's interest to cheat. My interpretation of the disagreement between Hobbes and the Fool follows Jean Hampton's interpretation. See Hampton, *Hobbes and the Social Contract Tradition*, 55–56.

the outcome where we both do our parts to the outcome where neither of us do our parts. So if I perform first, the other party will perform, and I will achieve my most preferred outcome.” Thus, it is rational for a party to perform first.

Hobbes is now in the position where contracts will be kept in the state of nature, which compromises his political argument that a sovereign is needed to punish those who violate their contracts. We now have a moral solution to the problems of the state of nature. Before pressing for political solutions, we should investigate the moral solution to assess its success and to see what problems remain for a political solution to address.

2.1.7 The Source of All Constraint: Mutual Advantage

We have now examined all of the sources and forms of constraint that Hobbes imposes on individuals: laws of nature and contractual obligation. This section will review these constraints, and try to bring the logic of the argument to the fore.

In the state of nature, the unlimited right of nature establishes a baseline against which all constraint must be justified. If one cannot justify, by some appropriate manner, a principle of constraint, then there will be nothing but the right of nature, and the state of nature will persist.

Against this baseline, Hobbes offers us two sources of constraint: the laws of nature and obligation. Hobbes justifies the laws of nature by showing that, given human nature, each individual must find these laws to be rational *viz.* in their interest, promoting their fundamental preferences. We obey the laws because obeying the laws will best preserve our lives, and we are structured in such a way that the preservation

of our life causes us pleasure, and the opposite causes pain.⁴⁰ These laws of nature are natural, not artificial, insofar as they apply to each human independent of any human choice or institution. Whatever choices we make, whatever institutions we create, and howsoever we may resolve to the contrary, we cannot avoid or undermine the laws of nature; no matter what, we must want things conducive to our continued life and happiness, and peace will be the most conducive means to that end. And the laws of nature are necessary and sufficient means to peace. So the laws of nature represent a kind of natural causal knowledge: obeying the laws is most conducive to peace, which is most conducive to the preservation of life and happiness. The reason that these laws are normative for us is that they connect with our fundamental interests: they show us the most effective means to what we desire. If we did not desire the effects to which they led us, then, even if they were true causal knowledge, they would provide us with no normative principles of action.

Obligation is in part justified by individual choice, and in part justified by the laws of nature. Obligation arises when an individual contractually renounces or transfers a right. Thus, unlike the laws of nature, obligation is primarily dependent on an individual's decisions: I become obligated only insofar as I choose to be, and to do only that which I choose to be obligated to do. Thus obligations are primarily artificial, not natural: what I am obligated to do is primarily determined by my decisions, not by natural facts.

Yet obligations are only *primarily* artificial: obligation is also partly rooted in the laws of nature in two ways. First, the third law of nature instructs one to perform covenants made (I.xv.1). For Hobbes, this is equivalent to claiming that performing

40. We could identify a 'false' claim advancing a law of nature insofar as it did not tend to our preservation, or to the fulfilling of any fundamental and universal (*viz.* shared by all humans) desire.

contracts made is conducive to peace, life, and happiness, and thus in the interest of each individual. Thus, while the content of an obligation is artificial, the requirement to perform an obligation is naturally grounded in the laws of nature. While, for the most part, I am obligated to do something only insofar as I *choose* to be obligated to do so, once I have made that choice and contracted to do so, I am obligated to perform my part.

The second sense in which obligations are naturally grounded shows that in some cases I am required to obligate myself.⁴¹ The second law of nature instructs one to contractually lay down some of one's natural right for the sake of peace (I.xiv.5). Thus, I am required to lay down some of my natural right for the sake of peace, even if I am disinclined to do so. Should I refuse an invitation to lay down some of my right in order to join a society with others, I would be acting contrary to the laws of nature. Members of that society could complain to me that my actions were contrary to the laws of nature, and thus were contrary to my own interests. They should be able to show this to me, and I should agree. If I cannot come to see that it is in my interest to join society, then I have made a mistake—one that I am likely to discover sooner rather than later insofar as this society now has an incentive to enslave or destroy me. Having revealed myself to be disinclined to restrict my behaviour for the sake of peace, I am now a threat to the peace of others.

What should now be obvious is that Hobbes justifies all constraints by showing them to be in our own interest. The laws of nature are natural causal knowledge, but they motivate us to act only insofar as they connect with our practical reason—with our weighing of the expected consequences of different courses of action, and selecting

41. "Required to obligate myself" may sound awkward, but note that it is worded in this way because of the technical meaning Hobbes reserves for "obligated." The laws of nature do not obligate one, though they do require one to act in certain ways. Obligations arise only from contracts.

that action most conducive to our preferences. We carry out our obligations only to the extent that carrying them out is in our interest: were it ever in our interest to violate a contract, we could and should do so. However, as the third law shows, Hobbes believes that such a situation could not arise—that, practically speaking, it will always actually be in our interest to keep our contracts. Given that, for the most part, we are free to choose to contract or not to contract, we will only form contracts that we believe to be in our interest. Thus all constraints on our behavior are derived from our prudential interests—from what our practical reason tells us will make our lives go best.

While Hobbes offers justifications for constraints that are relative to each individual, he is able generate one set of constraints for all people insofar as he focuses on what is common to all humans and relies on calculations of possibility and necessity.⁴² Hobbes achieves this combination by beginning with an account of human nature—of what all human individuals must share. According to Hobbes’s account, each individual wants that which is most pleasing and least displeasing. Things conducive to life must be pleasing, and things destructive of life must be displeasing, given the way that pleasure and pain are produced in humans. Hobbes acknowledges that human desires are variable, and thus focuses on those things that humans cannot help but seek or avoid. Thus, the laws require what humans must seek, and proscribe what humans must avoid. The reason that they are useful is that humans have tended to be mistaken about what actions will bring about the desired consequences. Hobbes thinks that wars have been the result of mistaken reasoning about how best to fulfill desires. And with respect to obligation, Hobbes deals with obligation *formally* for

42. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 90. For more on Gauthier’s analysis of Hobbes in this regard, see 2.1.5.

the most part—that is, without reference to the content of obligations, but only with reference to the structure of obligations. This allows Hobbes to discuss obligations affording for the provision of variable and specific desires without actually specifying or appealing to any non-universal desires. Thus, while Hobbes derives a set of constraints that are individually appealing—that I like from my point of view, and that you like from your point of view—he also derives a set of constraints that are generally or universally appealing—the laws of nature speak to all individuals insofar as they are human. But the laws are universal in another way too: not only do they appeal to desires that each individual has, they also provide a method for satisfying each individual’s desires to the greatest extent possible. As I said before, the precept “kill anyone who has some possession that would make your life go better, and take that thing” is capable of speaking to each individual insofar as it makes reference only to desires that all humans have, but it is a self-defeating strategy. The laws of nature coordinate our activities so that we may all achieve our ends. They do not just provide for individual advantage, they provide for *mutual* advantage. Not only is it rational for each of us to follow the laws of nature, it is rational for each of us to follow the laws of nature given that everyone else is following the laws too. The laws provide for the interests of each individual, and of all individuals simultaneously.

Thus, Hobbes justifies all constraints, whether they derive from the laws of nature or from obligation, by showing them to be mutually advantageous. Insofar as a proposed principle cannot provide anyone any advantage, or can provide advantage to some but only at the expense of others, it cannot represent a true constraint on individual behaviour.

2.2 Hume

Hume first articulates his theory of justice in *A Treatise of Human Nature*, published between 1739 and 1740, and re-articulates it in *An Inquiry Concerning the Principles of Morals*, first published in 1751.⁴³ As Barry notes, Hume seems to have presented the seeds of three distinct kinds of theory of justice: his circumstances of justice offer a mutual-advantage theory of justice akin to Hobbes's theory; his specification of the principles of justice offers a utilitarian approach to justice; his examination of why justice is a virtue, his account of sympathy and his associated impartial spectator device present a theory of justice and morality as impartiality.⁴⁴

This section focuses exclusively on Hume's definition of a convention and his account of "the circumstances of justice."⁴⁵ While Hume develops his own complete mutual advantage theory, it would be largely redundant to pour over the whole thing as many of the details are the same as in Hobbes. I focus on the circumstances of justice and the concept of a convention because Hobbes does not explicitly develop such a concepts, though they are highly relevant to mutual advantage theories. Moreover, the circumstances of justice provide us with a useful way to conceptualize the nature of this thesis: the thesis examines the principles of justice that arise when the circumstances of justice are attenuated in specific ways.

43. *A Treatise of Human Nature* is broken up into Book, Part, and Section, and references will be provided in the same form as those for Hobbes's *Leviathan* with book in upper-case Roman numerals, part in lower-case Roman numerals, and section in Arabic numerals. However, page numbers will also be provided because the sections are still rather long. *An Inquiry Concerning the Principles of Morals* is less thoroughly structured, so references for that work will be provided in the standard format of page numbers provided in footnotes.

44. Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice*, 146.

45. John Rawls, *Lectures on the History of Moral Philosophy*, ed. Barbara Herman (Cambridge, Mass.: Harvard University Press, 2000), 58.

2.2.1 Circumstances of Justice

Similar to Hobbes, Hume attempts to derive principles of justice from human nature and the circumstances in which humans find themselves. The interaction of these two factors is examined in Hume's discussion of what Rawls calls "the circumstances of justice."⁴⁶ Hume tries to explain the development of justice by reference to facts about human nature and the world:

[T]here are some virtues, that produce pleasure and approbation by means of an artifice or contrivance, which arises from the circumstances and necessity of mankind. Of this kind I assert *justice* to be.⁴⁷ (*Treatise* III.ii.1)

Rawls stresses that Hume "wants to show that morality and our practice of it are the expression of our nature, given our place in the world and our dependence on society."⁴⁸ For Hume, we develop principles of justice (1) because resources are only moderately scarce, (2) because humans have only limited benevolence, (3) because humans are of roughly equal strength and intelligence, and (4) because we are somewhat dependent on each other. Hume claims that justice is only intelligible given these circumstances, and shows this by considering cases where these conditions do not hold. Before we go on, however, it is important to note that Hume generally has a narrow understanding of justice, limited to rules governing property.⁴⁹ This narrow conception of justice is obvious in Hume's discussion of the following cases, where he slips back and forth between discussion of "justice" and "property" as though they

46. Rawls, *Lectures on the History of Moral Philosophy*, 58.

47. Hume, *A Treatise of Human Nature*, 477.

48. Rawls, *Lectures on the History of Moral Philosophy*, 51.

49. *ibid.*, 57.

were one and the same thing.

Moderate Scarcity

First, Hume considers a case where nature provides us with such an abundance of resources that each of us could satisfy even our most luxurious and frivolous desires with no effort or industry.⁵⁰ Luxury cars grow on trees, mansions grow from the ground with no tending, and food and wine are abundantly available. In this case, Hume claims, justice would not be invented because it would be useless: insofar as there is no need for any restraint, inventing restraints would serve no purpose.⁵¹ But likewise, justice would not arise in the opposite case of extreme scarcity, where no matter how much we labour to produce the necessities of life, most will die of deprivation and all will be miserable.⁵² Hume claims:

The *use* and *tendency* of virtue is to produce happiness and security, by preserving order in society. But where the society is ready to perish from extreme necessity, no greater evil can be dreaded from violence and injustice, and every man may now provide for himself by all the means which prudence can dictate or humanity permit.⁵³

In both cases of extreme scarcity and extreme abundance, there is no advantage to be gained from mutual restraint. Thus, Hume thinks that justice would only arise under conditions of moderate scarcity, where labour and industry are required to provide for our interests, but where they are also sufficient to provide for at least the basic interests of most of society.

50. Hume, *An Inquiry Concerning the Principles of Morals*, 15.

51. *ibid.*

52. *ibid.*, 17.

53. *ibid.*

Moderate Benevolence

Second, Hume considers the case where each human is extremely generous, and is just as concerned about the well-being of each other person as she is about her own well-being. In this case, justice would be useless: it would not be able to add anything to anyone's well-being, and so would not be invented.⁵⁴ Likewise, however, in a society of "ruffians" who had no concern for the well-being of others Hume thinks justice would not arise. Hume's argument is difficult to reconstruct because he simply considers the case of a virtuous man dropped into such a society, and notes that justice would be of no use to this man.⁵⁵ And, as Barry notes, a deficiency of generosity should pose no impediment to Hume's principles of justice insofar as those principles are rationally demonstrated to be in each individual's interest:

[I]t would make no difference as far as justice is concerned if people were totally selfish. The circumstances of justice would still obtain. Justice, Hume says, is founded on self-interest: 'itself alone restrains it.' The question is not one of the 'wickedness or goodness of human nature' but of 'the degrees of man's sagacity or folly.'⁵⁶

Thus, even in a society of completely ungenerous misers, justice would have a role to play in helping the members of such a society get along better together, and each individual would be able to recognize and follow the principles of justice so long as such individuals were still capable of acting according to what they perceived to be in their best interest.

54. Hume, *An Inquiry Concerning the Principles of Morals*, 16.

55. *ibid.*, 18.

56. Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice*, 158.

There are multiple ways of interpreting the opposite extreme to full benevolence, each of which seems to be roughly accurate. First, Barry thinks that we should interpret the opposite extreme to full benevolence not to be one of extreme meanness, rather one of “*unintelligent* self-interest.”⁵⁷ Here, there are a number of ways we could be unintelligent that would prevent the development of justice. First, we could be bad at determining what was in our interest. Our reasoning abilities could be quite short-sighted, or very easily overwhelmed by strong passions, so that whenever we had a passing desire, our reason immediately identified that desire to be what determines what we ought rationally to do. So, when I feel the urge to steal from someone, I immediately reason that stealing is the wise thing to do for the sake of making my life go better. If we were that way, we would consistently find, after the fact, that what we had identified initially to be rational was not so.

Second, we could be extremely weak willed. On this interpretation, we have sufficient ability to determine what course of action is in our interest, yet this determination is rarely, if ever, capable of motivating us to carry out that course of action. Instead, our passions have greater control over our actions. We would be incapable of observing any constraints. In this manner, I could be well aware that stealing from some fellow is definitely unwise, yet feel so compelled to do so by my desires that, in the face of reason, I steal.

Barry also offers us a third interpretation of the opposite extreme to full benevolence, according to which humans find it more important to be able to hurt others than to provide for their own security.⁵⁸ On this interpretation, humans are not supposed to be deficient in determining their interests, nor are they too weak-willed

57. Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice*, 158.

58. *ibid.*

to act accordingly; rather, their interests are roughly ‘evil’ in that they would rather cause harm to others than be secure and comfortable in their own lives. Such interests would preclude any cooperation, and thus preclude the development of principles of justice.

Thus, the appropriate way of understanding the moderate benevolence condition is not that we are moderately self-regarding and moderately other-regarding—that is, that we are concerned with our own welfare but also concerned with the welfare of others—rather, it is that we are neither fully benevolent, nor evil. Our preferences and reason are not set up so that we would naturally decide to make everyone else’s lives go best at our own expense, in which case there would be no conflicts of interest for justice to resolve, nor are they set up so that we would naturally decide to make everyone else’s lives go badly even at our own expense, in which case conflicts of interest would be insurmountable. Rather, our preferences and reason are such that our interests conflict because we will naturally promote our own interests at others’ expense, but those conflicts are resolvable both because our preferences are such that we would rather make ourselves better off than make others worse off, and because we are capable of identifying and acting on mutually advantageous principles constraining our actions. We are not fully benevolent, but we are neither unintelligent, nor weak-willed, nor evil.

Rough Equality

Hume then considers what would happen if there was some inferior species living among humans.

Were there a species of creatures intermingled with men which, though

rational, were possessed of such inferior strength, both of body and mind, that they were incapable of all resistance and could never, upon the highest provocation, make us feel the effects of their resentment, the necessary consequence, I think, is that we should be bound, by the laws of humanity, to give gentle usage to these creatures, but should not, properly speaking, lie under any restraint of justice with regard to them, nor could they possess any right or property exclusive of such arbitrary lords. Our intercourse with them could not be called society, which supposes a degree of equality, but absolute command on the one side, and servile obedience on the other.⁵⁹

Hume thinks that in such circumstances, justice between humans and these creatures would be useless to humans because anything we wanted from those creatures we could take without their consent and without any cost to ourselves (in terms of retribution exacted by those creatures).⁶⁰ The inferior species could gain from our restraint, but we could only lose. Thus justice would not arise. Hume thinks that this captures the circumstance of animals.⁶¹ So justice, insofar as it develops, will only develop to restrain rough-equals—those who are able to “make us feel the effects of their resentment.”⁶²

Mutual Dependence

Finally, Hume considers a case where each human being is completely independent of all others because each is wholly self-sufficient.

59. Hume, *An Inquiry Concerning the Principles of Morals*, 21–22.

60. *ibid.*, 22.

61. *ibid.*

62. *ibid.*, 21.

Were the human species so framed by nature as that each individual possessed within himself every faculty requisite both for his own preservation and for the propagation of his kind, were all society and intercourse cut off between man and man...it seems evident that so solitary a being would be as much incapable of justice as of social discourse and conversation.⁶³

This case is similar to the case of natural abundance: in the latter we are independent because nature offers all we could need, given our constitution and our desires, while in the former we are independent because our constitution and desires are such that, given what nature has to offer, we are completely self-sufficient. The significant difference between this case and the case of natural abundance is that in this case we are independent from each other: other humans have nothing to offer that we need. The focus in this case is independence from other humans. Here, justice could not arise because it would be useless. However, Hume describes a family society (which is, of course, the kind of society we find ourselves in) where two members of opposite sexes must come together to produce offspring. This condition on reproduction leads to the creation of families, with rules governing behaviour within the family. Families could get together to form small cooperative societies, and the rules would expand to govern members of that group but only members of that group. But then a number of these small societies could start to cooperate, etc. In such a society where individuals are moderately dependent on each other, justice arises.⁶⁴ The strength of the bonds of justice will be proportionate to the level of interdependence.⁶⁵

63. Hume, *An Inquiry Concerning the Principles of Morals*, 22.

64. *ibid.*, 23.

65. *ibid.*, 35–36.

Instability of Possession

To these four features of the “circumstances of justice,” a fifth can be added. In *A Treatise of Human Nature*, Hume notes that we are secure in enjoying the inward pleasures we feel, for no one can take these from us, and that our bodies are open only to destruction—another person can attack and injure my body, but cannot appropriate it to benefit from it directly. However, our “external goods,” the possessions we enjoy and depend upon, can be taken from us that others may enjoy them. The fact that, unlike our pleasures and our bodies, our external goods can be taken from us, and can then provide whatever enjoyment or service, which they would have provided to us, to whoever took them, undermines the stability of our holdings. This is an obstacle to society, and is one of the conditions for the development of justice.⁶⁶(*Treatise* III.ii.2). We can construct a case as Hume did above: if we imagined that all external goods ceased to exist as soon as they were removed contrary to the will of their possessors, presumably justice would not arise, since Hume’s conception of justice is mostly confined to property rights.

Thus, Hume thinks that the moderate scarcity of nature’s resources, our moderate benevolence, our rough equality of strength and intelligence, our mutual dependence, and the instability of our possessions are individually necessary and jointly sufficient for the creation of justice. If any of these conditions were lacking, we would not have invented principles of justice. But given that these conditions obtain, we continue to recognize principles of justice. These circumstances define the problems that principles of justice are intended to resolve.

66. Hume, *A Treatise of Human Nature*, 487–489, 494.

Circumstances of Justice Are a Matter of Degree

It is important to note that the circumstances of justice, as Hume explains them, are to a great extent a matter of degree. Principles of justice will be useful, and hence will emerge, only where resources are moderately scarce, where individuals are not too benevolent nor too malevolent, where individuals are roughly equal, and where individuals are somewhat dependent on one another. Each of these conditions picks out a range on a spectrum of conditions that is suitable for justice; each condition can hold to a greater or lesser degree. Moreover, the limits of each range are likely to be unclear: exactly where we draw the line between sufficiently benevolent individuals and individuals who are too malevolent for justice is likely not only to be a contentious matter but one where there is no clear answer to be had. One way to characterize the project at hand is as an exploration of the fringe limits of the circumstances of justice: How weakly can the circumstances of justice hold—in particular, that of rough equality—and yet moral principles still arise?

2.2.2 Convention

Hume claims that the only way to remedy the problems posed by the circumstances of justice discussed above is to enter into convention.

This can be done after no other manner, than by a convention enter'd into by all the members of society to bestow stability on the possession of those external goods, and leave everyone in the peaceful enjoyment of what he may acquire by his fortune and industry.⁶⁷(*Treatise* III.ii.2)

67. Hume, *A Treatise of Human Nature*, 489.

A convention is not a promise.⁶⁸ Rather, “it is only a general sense of common interest; which sense all the members of the society express to one another, and which induces them to regulate their conduct by certain rules.”⁶⁹ Rawls finds that a convention has two features, and that three conditions must be met for a convention to hold. The two features of a convention are its rules and a “shared awareness of a common interest, which all members of society express to one another, and the recognition of this common interest leads all to regulate their conduct by the rule(s).”⁷⁰ The three requirements that must be met for a convention to hold are (1) that the common interest (understood as the interest of each member) is expressed and recognized by the members of society, (2) that there are rules that are sufficiently straight-forward, intuitive, and publicly recognized to bring about that which is in the common interest, and (3) that it is reasonable for each member to predict that most members will follow the rules most of the time.⁷¹ Rawls notes that, although a convention is not a promise, it shares with promises a kind of mutuality or reciprocity: each member follows the rules only insofar as others do too, thus each member’s behaviour is conditional on the behaviour of others in the same way as we find in promises.⁷²

Hume thinks that the content of the convention will be limited to rules governing property, since this will be sufficient to ensure “perfect harmony and concord.”⁷³ In small societies, this convention will be self-enforcing insofar as the common interest is apparent and repeated interactions provide reasonable confidence that others are respecting the rules. In larger communities, the temptation to cheat, combined with our

68. Hume, *An Inquiry Concerning the Principles of Morals*, 122; Hume, *A Treatise of Human Nature*, 490.

69. *ibid.*

70. Rawls, *Lectures on the History of Moral Philosophy*, 60.

71. *ibid.*, 60–61.

72. *ibid.*, 60.

73. Hume, *A Treatise of Human Nature*, 491.

tendency to make mistakes about what is in our interest and thereby fail to recognize the common interest, leads to both the moral emotion of justice through sympathy and to the creation of ‘political societies’ designed to enforce the advantageous rules of the convention.⁷⁴

2.3 Those Excluded From Moral Rights

Now that the fundamentals of Hobbes’s and Hume’s mutual advantage contract theories have been explained, we can examine the reason why they believed animals and human marginal cases to be denied moral rights. This section will examine the extent to which animals, disabled humans, human infants, and elderly humans *seem* to be excluded from consideration according to Hobbes and Hume. The point of this section is to see why both critics and advocates of mutual advantage social contract theory have taken it to exclude the groups in question. In chapter 3 I will revisit these purported exclusions and argue that, in fact, mutual advantage contractarianism includes the majority of these cases.

Animals

First, Hobbes explicitly rules out the possibility of obligating oneself to animals because animals lack the capacity to form contracts (I.xiv.22, II.xxvi.12). Animals lack the capacity for the kind of language needed to form contracts; they can neither comprehend when one promises to perform a certain action, nor can they promise to reciprocate (I.xiv.22, II.xxvi.12).⁷⁵ So obligation is ruled out as a source of moral

⁷⁴ Hume, *A Treatise of Human Nature*, 498–499, 534–539; Hume, *An Inquiry Concerning the Principles of Morals*, 34–35; Rawls, *Lectures on the History of Moral Philosophy*, 63, 67–68.

⁷⁵ La Barbera, *States of Nature: Animality and the Polis*, 50.

constraint regarding animals. Thus, the exclusion of animals from obligation relies on two premises for Hobbes: first, that animals and humans cannot communicate in the manner needed for covenants, and second, that expressed covenants are required for obligations to arise.

Hobbes also claims that animals lack foresight, concerned only with the present moment, and because of this deficit in foresight they also lack causal knowledge (I.xi.4).⁷⁶ While Hobbes did not explicitly connect this claim with the exclusion of animals from obligation, we can see that a deficit of foresight and causal knowledge would bolster the exclusion of animals. For even if animals were granted the power to communicate with humans, should they have little concern for the future and should they lack knowledge of the effects that stem from different courses of action, they might still lack the ability to covenant.

Given that obligation is ruled out, the only source of moral constraint left is the law of nature. Hobbes makes no explicit mention of animals when identifying his seventeen laws of nature, yet there is no reason for thinking that the list of laws that Hobbes identifies is meant to be exhaustive. We are thus left to seek an explanation for why animals are excluded from the protection of the laws of nature elsewhere.⁷⁷

We find in Hume an argument that explains why animals would be excluded from the laws of nature. Unlike Hobbes, Hume argues that animals are capable of causal reasoning and of communication.⁷⁸ For Hume, animals are excluded from the protection of morality because their relative weakness makes it such that any principle

76. Hobbes may contradict himself on this matter. See his discussion of animal prudence at I.iii.9.

77. Hobbes has passages that explain why animals are not duty-bound by laws, and though in these passages he is talking about legal laws, not natural laws, the reasoning would apply equally to the laws of nature. Because animals cannot understand the laws, they cannot be duty-bound by them (II.xxvi.12 and II.xxvii.23). Yet this does not explain why animals are precluded from being protected by the laws of nature.

78. II.ii.12 and I.iii.16 in Hume, *A Treatise of Human Nature*, 226–229.

calling for human restraint toward animals would be disadvantageous to humans.⁷⁹

In general, any principle of action could be a law of nature so long as (a) the principle offers each and every individual a net improvement in some fundamental value shared by all, like the value to an individual of continuing to live, and (b) there is no alternate principle that would offer a greater gain in terms of that value. If anyone is to be excluded from the benefit of such laws, it should not be because Hobbes did not explicitly include them in the laws, but rather because including them in the laws would violate one of these conditions.

Animals are excluded from the protection of the laws of nature because their inclusion will not offer all individuals an improvement in some value that all humans possess. First, including animals would impose a cost on humans. Depending on the kinds of consideration afforded to animals, farmers could be disallowed from forcing animals to breed, raising them in captivity, or killing them for resources. Humans could be prevented to some extent from developing in animals' natural habitats. In general, there would be a host of activities from which humans benefit and animals lose that would be proscribed by laws of nature protecting animals.

This cost to humans could be overridden if there was a greater benefit to be reaped from such constraint. In order for such constraint to be justified, the benefit must (a) accrue to each individual, and be greater for each individual than both (b) the cost of this constraint and (c) the potential benefit from any other potential law. Hobbes confines himself to justifying principles in terms of benefits to shared values because, if any individual lacked the value in question, such an individual would have no reason to obey the proposed law. For instance, if someone proposed that we

⁷⁹ For more discussion of Hume's reasoning on this matter, see section 2.2.1 on page 62. Hume, *An Inquiry Concerning the Principles of Morals*, 21–22.

should never deliberately destroy a tree, because trees are very beautiful in a number of ways—the appearance and smell of their fiery leaves in autumn, their shade on a hot summer’s day, their sound on windy days, etc.—it could be the case that such a principle maximized some people’s expected advantage. Such people recognize that there is a cost to such constraint, but find that constraint to be quite small, and to be easily outweighed by all of the things they like about trees not being destroyed. The problem is that many people likely lack the preference structure that would make the benefits of not destroying trees outweigh the costs of such constraint. Such people would find no reason to obey such a law of nature, and thus could not be expected to comply. Thus, if we are all to obey the laws of nature, which Hobbes takes for granted as a desideratum of laws of nature, then such laws must be justifiable to each individual on terms that individual endorses. This is what is meant by the first condition: each person must consider herself to benefit from the proposed constraint.

Yet the case of animals is likely to be like the case of trees. Because Hobbes thinks that animals cannot appreciate the consequences of their actions, take an interest in their futures, or bind themselves in contract, it is clearer for Hobbes that animals can offer us no benefit in response for our restraint. For Hume, animals are so weak that they cannot resist humans and thus any benefit we can get from them we can get more efficiently by force than by constraint.⁸⁰ The only remaining benefits to be derived come from our own preferences concerning the treatment of animals. Though it is certain that a greater number of people would prefer it if no animals were killed than would prefer it if no trees were cut down, it is surely not currently the case that *everyone* wants animals not to be killed. Thus, arguably, the first condition is not met, and has never been met in human history.

80. Hume, *An Inquiry Concerning the Principles of Morals*, 21–22.

The second and third conditions are even less likely to hold true than the first condition. There is likely some subset of the people who find that they benefit from the not killing of animals who would nevertheless find that they stood to benefit more from a principle allowing some killing of animals, and there is likely some subset of that subset who would also find that the benefits derived from the not killing of animals did not outweigh the costs of such constraint. So even if some people would appreciate a principle proscribing the killing of animals, fewer of them would find an alternate principle allowing for some killing of animals was not better for them, and fewer still would find that the benefits of such constraint outweighed the costs. Thus, such a principle could not be a law of nature; it could not offer each individual a reason to constrain his or her behaviour. This is the traditional reason that animals are considered to be excluded from moral protection by mutual advantage social contract theory.

Children

Hobbes explicitly excludes human children from the laws of nature: because children lack reason and language, they are not bound by the laws of nature (I.v.18, I.xvi.10, II.xxvi.12). However, Hobbes wavers on whether or not children are thereby excluded from covenants; as we would expect, he claims:

Over natural fools, children, or madmen there is no law, no more than over brute beasts; nor are they capable of the title of just or unjust, because they had never power to make any covenant or to understand the consequences thereof. . . . (II.xxvi.12)

Yet, Hobbes also claims that parents are the sovereign masters of their children, and that this sovereignty comes not from the fact that the parents created their children but rather from a contract between children and parents.

The right of dominion by generation is that which the parent hath over his children, and is called PATERNAL. And is not so derived from the generation as if therefore the parent had dominion over his child because he begat him, but from the child's consent, either by express or by other sufficient arguments declared. (II.xx.4)

This contract between children and parents is formed based on the immense power that the parents have over the children, "being able to destroy them if they refuse" (II.xvii.15). This obligation of children to parents extends beyond the period in which the children are completely dependent on the parents in recognition of the benefits that the children have received from their parents (II.xxx.11.).

Hobbes's exclusion of children based on their lack of reason and language fits best with the rest of what Hobbes has said about excluding animals from contracts. His argument for parents' dominion over children fits better with Hume's notion of a convention, and with his inequality-based arguments for exclusion from moral protection. The idea is essentially that because children are so weak compared to parents, parents are under no obligations toward them, and the children *must* do whatever the parents demand, where "must" means simply "will obey or face severe punishment." This argument is not restricted to parents: children stand in this relation to *all* adults.

Thus, children are excluded from all moral protection due to their lack of reason, lack of language, and relative weakness. These qualities define what it means to be

a “child” for the purpose of moral standing: once one can reason, speak, and resist, one earns the same moral status as all able human adults.

Remaining Human Marginal Cases: Disability, Illness, and Infirmary

This section will address the remaining human marginal cases, including human disability, illness, infirmity, and generally any other condition that causes a human individual to be in a comparable position to these. Examining the reasoning for excluding such cases will also make clear what features define the excluded positions. It should be emphasized that we are all likely to be “marginal” cases at various points in our lives.

Hobbes’s explicit discussion of disability, and of marginal cases more generally, is quite limited. “Natural fools,” “fools,” and “madmen” are excluded from the laws of nature and from covenants because they cannot understand laws or covenants, or author any actions (I.xvi.10, II.xxvi.12, II.xxvii.23). If one lacks the ability to reason, to understand, to communicate, or to covenant, one cannot be protected by obligation and one cannot be subject to the laws of nature. And for the same reason that animals cannot be protected by the laws of nature due to their relative weakness, one cannot be protected by the laws.

But certainly most marginal cases would not count as “fools” or “madmen” on this account. Many marginal cases, for instance people with learning disabilities or limited mobility, are perfectly capable of reasoning, understanding, communicating, and covenanting. So while marginal cases who are incapable of such cognitive and communicative acts are excluded from moral consideration by Hobbes, many marginal cases will be included.

However, even completely rational and vocal marginal cases may be excluded from the protection of the laws of nature and obligation based on the same reasoning that applied to animals: if some marginal cases are so weak, in body or mind, that they could not “make us feel their resistance and could never, upon the highest provocation, make us feel the effects of their resentment,” then they are too weak to be protected by the laws of nature or by obligation.⁸¹

Thus we get mixed results from mutual advantage contract theory with respect to disability, illness, and infirmity: So long as one is capable, by colluding with others or by a clever plot, of killing another human or making one’s resistance and resentment felt, one is protected by the laws of nature and by obligation. Otherwise, one is excluded from all moral protection.⁸²

2.4 Conclusion

This chapter has presented the fundamental elements of mutual advantage social contract theory, as articulated by Hobbes and Hume. This chapter also examined the reasoning that led Hobbes and Hume to think that animals and many marginal cases were excluded from moral consideration according to a mutual advantage theory. Chapter 3 will build on and revise these elements to articulate the most plausible mutual advantage contract theory of rights for animals and marginal cases. It will also articulate a stronger account of rights and protections for animals and marginal cases.

81. Hume, *An Inquiry Concerning the Principles of Morals*, 21.

82. Hume speaks of “the laws of humanity, to give gentle usage” to creatures who lack the strength to qualify for rights, and while this claim may fit well with the non-contractarian aspects of Hume’s thought, it does not fit well within a mutual advantage framework. *ibid.*

Chapter 3

The Contractarian Theory of Animal Rights

Chapter 2 examined the fundamental elements of mutual advantage contract theory as developed by Hobbes and Hume, and examined the reasons why contract theory *seemed* to exclude animals and marginal cases from moral consideration. This chapter answers the question, “What rights would animals and marginal cases receive from the most plausible social contract theory rooted in mutual advantage?” Section 3.1 revisits the fundamental elements of mutual advantage social contract theory discussed in chapter 2, and revises and bolsters them to yield the most plausible social contract theory. It then examines the precise requirements that must be met to qualify for rights or “protectorate status” according to contract theory.¹ Sections 3.2 and 3.3 deduce, from that contract theory, the rights and protections that animals and marginal cases should receive.² Finally, section 3.4 considers and refutes two objections to contractarianism that are common in the animal rights literature. Thus, this chapter

1. Protectorate status is explained in section 3.1.7 on page 101.

2. A substantial portion of this chapter—from section 3.1.5 to the end—has been published in *Les ateliers de l'éthique / The ethics forum*, and has been slightly modified and represented here with the permission of the publisher. See Matthew Taylor, “Grounding Animal Rights in Mutual Advantage Contractarianism,” *Les ateliers de l'éthique / The Ethics Forum* 9, no. 3 (2014): 184–207, doi:10.7202/1029065ar.

hopes to articulate the most plausible contractarian account of rights for animals and marginal cases, and to show that—even if it remains counterintuitive—the theory is consistent and deserves to be taken seriously. Chapter 4 will then present the dominant theories of animal rights, and chapter 5 will argue that the mutual advantage contract theory offers the best account of rights from a security perspective.

3.1 Contractarianism Revisited

3.1.1 Human Nature and Rights

Hobbesian mutual advantage social contract theory has developed considerably since Hobbes and Hume. Most notably, David Gauthier's *Morals by Agreement* offered an advanced articulation of a Hobbesian moral theory. But because Gauthier's theory cannot generate rights, it is relatively unhelpful to this thesis's task of generating a social contract theory of rights. This section explains why this thesis starts with Hobbes's ideas rather than Gauthier's.

In *Morals by Agreement*, Gauthier articulates a moral theory inspired by Hobbes and Hume. His aim is to show that it is rational to observe moral constraints on one's actions.³ To do so, Gauthier offers two central converging arguments. The first, concerning what is rational, demonstrates that rational utility-maximizers (a) would observe the principle of minimax relative concession in bargaining to establish society, (b) would thus adopt minimax relative concession as the principle governing society, and (c) would observe the principle of minimax relative concession in co-operative interaction.⁴ The second argument, concerning what is moral, claims that

3. Gauthier, *Morals by Agreement*, 4–5.

4. *ibid.*, 113–156, especially 145.

the principle of minimax relative concession would be chosen by an ideal actor from the Archimedean point (similar to Rawls's original position, but differing in details). Since any choice from the Archimedean point must be impartial by virtue of the specification of the choice situation, and since Gauthier takes impartiality to be the defining feature of morality, the principles chosen from such a perspective are moral principles.⁵ These two arguments, taken together, show that it is both rational and moral to observe the principle of minimax relative concession in social interactions.

Gauthier's social contract argument produces a theory of distributive morality: his contract theory yields the "principle of minimax relative concession," a moral principle determining how we will divide the surplus generated by our cooperative endeavours.⁶ But Gauthier is quite clear that his contract theory does not yield an account of moral rights.⁷ For moral rights like the right to life, and the correlative duty to not kill, Gauthier provides us with a Lockean account that must antecede the social contract.⁸ If society did not come to recognize natural rights, Gauthier's contract model would not apply and thus his ethical distributive principle would not be adopted. This leads me to assert that, in fact, Gauthier does not offer a social contract theory of rights. Why would he do this?

One strength of Gauthier's account, regarded as an improvement over Hobbes's theory, is that it appeals to a much thinner account of human nature. By a theory of human nature, I mean a set of claims about properties that are shared by all humans. Granting that nature produces a few exceptions, a theory of human nature will claim that nearly all humans share certain properties by virtue of being human, and that

5. Gauthier, *Morals by Agreement*, 233–267.

6. *ibid.*, 137–140, 145.

7. *ibid.*, 222.

8. *ibid.*, 190–232.

any individuals who lack those properties appear to be unhealthy or defective.⁹ Thus, the greater the number of properties included in an account of human nature, the more likely it is that the account of nature is false: it can never be the case that some assertion P is less likely to be true than the conjunction $P \ \& \ Q$. Not only do we run a greater risk of being wrong when we rely on a thicker account of human nature when building our moral theories, but we also run a greater risk of thereby failing to persuade those exceptional humans that they should endorse the principles that our moral theory advances. Thus the concern is not just with getting it wrong but also with failing to bring about as much compliance and cooperation as we otherwise could—a failure which can involve significant costs.¹⁰ If a handful of people remain unconvinced because our assumptions about human nature excluded them from our moral argument, and we happened across one such person in a dark deserted alley, we could be in real trouble. So, other things being equal (which they seldom are), a moral theory that appeals to a thinner account of human nature is preferable to one that appeals to a thicker account of human nature.

Gauthier's view appeals to quite a thin account of human nature: humans desire various things, and care that their desires should be met as best as possible. Thus, humans act so as to maximize the fulfillment of their most important desires.¹¹ By contrast, Hobbes adds to this account that (healthy) humans desire the continuation of their own lives, and do not desire to harm others—at least, not if harming others is not instrumental to the fulfillment of some other desire, as it might be when we

9. But see the concerns with appeals to health discussed in the footnote 11 in section 2.1.2 on page 24.

10. This cost will be examined in greater detail in chapter 5.

11. Gauthier, *Morals by Agreement*, 133–134, 143, 330–355.

punish those who cheat us.¹² Thus, by appealing to a thinner account of human nature, Gauthier offers a moral theory that is, other things equal, more appealing than that of Hobbes.

But other things are not equal: Hobbes's moral theory is able to justify a duty not to kill others, except in specific exceptional circumstances, while Gauthier's social contract theory is incapable of generating any such prescription.¹³ In theory, Gauthier's social contract theory is limited to the domain of distributing the gains of cooperation.¹⁴ Gauthier's social contract theory could offer a procedure for generating a complete moral theory: parties to the contract endorse the minimax relative concession rule as a principle governing their bargaining.¹⁵ The parties thus consider the utility (viz., the extent to which their highest desires are satisfied) they would get from cooperating in accordance with various assignments of moral rights and duties. They might start by making a maximal claim: perhaps they claim all rights for themselves, while denying rights to anyone else. But this is likely overreaching: that would be equivalent to claiming more than one contributes, creating an incentive for others to exclude you from the cooperative endeavour altogether.¹⁶ Supposing we settled the question of the maximal claim, agents would then accept concessions in accordance with the minimax relative concession rule until an agreement had been reached that afforded each party a gain over the non-cooperative state of nature. The process is sufficiently clear; however, with only the process and the thin account of human nature as theoretical resources, we cannot deduce which assignment of moral

12. See sections I.vi.1 to I.vi.12, and I.vi.47 in Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, 27–30, 32.

13. See Hobbes's discussion of the "laws of nature," especially the first and second laws. *ibid.*, I.xiv.4–I.xiv.5, I.xv.1, I.xv.16–I.xv.33; pp. 80, 89, 95–98.

14. Gauthier, *Morals by Agreement*, 133–134.

15. *ibid.*, 145.

16. *ibid.*, 133–134.

rights and duties would be selected. To answer this question theoretically, we would need more information about the preferences of the parties to the contract. Perhaps the question could be settled practically by getting together and running through the process, but then we would be left with a social contract, not a social contract theory, and an actual social contract is not as helpful as a social contract theory. The idea that we could actually form a contract in practice that would be binding on new entrants to society is terribly implausible. The real strength of social contract theory lies not in the idea that we actually agree, at a point in time, to a set of principles, but rather in the idea that at all times we all ought to agree to and obey a specific set of principles. Gauthier's social contract theory cannot give us such an account of the assignment of rights and duties.

This is quite a serious weakness of Gauthier's social contract theory. It fails to offer us a set of rights and duties, offering us instead a principle for dividing the benefits we create by cooperating with one another. While that principle of distributive morality is useful, its usefulness requires that a set of moral rights and duties has already been recognized and adopted. If we are still killing and stealing from one another, the minimax relative concession rule will not help us much. Thus, I contend that the assignment of moral rights and duties is a more important task than the specification of a distributive principle for dividing the gains of cooperation: the benefits of fulfilling the last task will be nil until the first task is completed, while the benefits of the first task would be substantial even if the last task had yet to be completed. And these benefits justify appealing to a thicker conception of human nature. Moreover, since the goal of my project is to determine whether or not social contract theory can justify animal rights, I will need to turn my attention away from Gauthier and back toward

the more traditional contractarian theories that aim to justify a set of rights.

3.1.2 Human Nature, Misanthropic Desires, and Coalition Games

What is the thinnest account of human nature capable of yielding a theoretical assignment of moral rights and duties? Returning to Hobbes and Hume would be a good way to start this search. Hobbes and Hume add to Gauthier's account the claims that (a) humans want to keep living, and (b) humans do not want to hurt others except as a means to the fulfillment of other desires.¹⁷ Are these claims reasonable? Would they lead to exclusions from our moral theory that we should be worried about?

Both assumptions should strike us as false in their current universal formulation. There are cases where a human may prefer to die, and there are certainly humans who desire to harm other humans non-instrumentally. Is there a way to qualify the two claims such that they can still play a role in our moral theory without being so patently false? We could modify the first claim to assert that most humans for most of their lives prefer their continued living to their dying, and even when they cease to have this preference for life over death, they always prefer not to be killed by someone else without consenting to such killing. Arguably, in most cases where someone voluntarily ends her or his own life, she or he desires to keep living, but fulfilling that desire would prevent the fulfillment of another desire that she or he considers more important. Consider a soldier who jumps on a grenade to protect her comrades in arms, or a father who works his way to an early grave to feed his children. The father and the soldier both want to keep living, but another desire is more important to them—to protect her comrades and to provide for his children.

17. Hobbes, *Leviathan: With Selected Variants from the Latin Edition of 1668*, I.vi.1–I.vi.12, I.vi.47; pp. 27–30, 32.

Had there been another way to fulfill the latter desires that did not involve sacrificing their lives, the soldier and the father would have preferred and chosen that way. And in more typical cases of suicide, it is still plausible to think that a desire for life persists. Someone in the early stages of a progressive and debilitating illness who decides to end his or her life likely still has a desire to keep living, but more strongly desires to be free of pain, or to maintain independence, a unified sense of self, or autonomy. So the claim that humans almost always desire to keep living sounds rather plausible, and the rider that, even when that desire disappears or gets outweighed, people always desire not to be killed without their consent sounds at least equally plausible. If so, then we have already taken a good first step toward ensuring a right to life; however, the case of misanthropes still stands in the way.

What do we say about humans who want to harm other humans non-instrumentally?¹⁸ What has historically been said will not do. Hobbes and Hume both assumed that humans do not want to harm other humans, except when doing so contributes instrumentally to the actor's fundamental interests, like the interest in self-preservation. Or at least, if we do want to harm other humans non-instrumentally, this desire is less important to us than our desire for our own continued welfare. That may certainly be true of the vast majority of people, but what do we say about the exceptions? There

18. Vallentyne has argued quite forcefully that contractarian theories ought to take people's preferences as they are, and cannot appeal to counterfactual preferences. He has further argued that the worst case scenario for mutual advantage contractarianism is not mutual disinterest, but rather that some should desire harm for others. Moreover, those contractarian theories that reject appeal to moral intuitions cannot exclude these misanthropic preferences as morally irrelevant or morally wrong. Thus, I focus on the worst-case scenario Vallentyne has identified and attempt to show that contractarians can set such preferences aside to some extent without needing to appeal to antecedent moral intuitions. Peter Vallentyne, "Contractarianism and the Assumption of Mutual Unconcern," in *Contractarianism and Rational Choice: Essays on David Gauthier's Morals By Agreement*, ed. Peter Vallentyne (Cambridge: Cambridge University Press, 1991), 71–75.

are certainly some peculiar individuals who greatly desire to harm other people because they seem to get some joy out of doing so. Moreover, some such individuals are willing to jeopardize their lives for the sake of fulfilling these misanthropic desires. Such individuals will not find Hobbes's account of morality persuasive: insofar as Hobbes's assumptions about the nature of the bargaining parties exclude these individuals, his theory will not apply to them. Should we revise Hobbes's assumptions about the nature of the bargaining parties to include these misanthropes?

Narveson argues that we should not adjust our assumptions to include misanthropes. He says that a theory of justice should aim to find principles on which the most inclusive group of people can cooperate. Misanthropes, Narveson thinks, are people with whom cooperation on any terms is impossible.

Let us suppose that morality is a kind of club—the “morality club.” Anyone can join—no problem. Those who join have certain responsibilities and certain rights, and we, the people who run this club, offer a package that we think no remotely reasonable person could really refuse; but nevertheless, some might. All we are saying is that *our* package is such that it must appeal to the widest set of people any set of principles *could* appeal to. Anyone who doesn't buy our package wouldn't buy any package compatible with living among his fellows on terms that they could possibly accept. If we can make good on this offer, then the objection that our morality is not, after all, truly universal is hollow. . . . It's universal in the sense of being as nearly universal as any set of restrictive rules could be. And that, we can argue, is universal *enough*.¹⁹

19. Narveson, *The Libertarian Idea*, 135.

While I originally found this argument lacking and sought for a better solution, I have been unable to find one, and have been led to conclude that Narveson's response is the best position available to contractarians.²⁰ But it does call for some important clarification: in particular, what does it mean for a "package" to be "compatible with living among his fellows on terms that they could possibly accept"?²¹

The contractarian logic on this issue is already familiar, and we find one instance of it in Gauthier's definition of a "rational claim." In developing a contractarian solution to the problem of dividing the benefits of cooperation, Gauthier envisions a bargaining situation where each party first makes a claim to a portion of the surplus, and then parties make concessions until they reach a mutually advantageous agreement. In determining what would be a rational claim, Gauthier notes that each party wants as much of the surplus as possible and so makes a maximal claim, but that this claim must be bounded by the interest each party has (a) in reaching an agreement (otherwise, no cooperative surplus will be produced), and (b) in not being excluded from cooperation.

Each person wants to get as much as possible; each therefore claims as

20. In earlier versions of this thesis I attempted to divide contractarian tools into "moral" and "political," arguing that rights and insurance-based protections were moral solutions, whereas threat-based protections were political. The idea was that moral solutions appealed to reasoned arguments that applied universally to all duty-holders, and thus depended less on detection and punishment mechanisms than threat-based political solutions, which involved a subset of duty-holders threatening the rest with punishment should they violate their duties. What I failed to notice was that the problem of misanthropic desires remained just as large a threat to moral solutions, effectively ensuring that there would be virtually no moral rights at all. Should there be one misanthrope who finds a moral right to life disadvantageous, the right to life would be a political rather than a moral right. Not only was this strategy unsuccessful, it was entirely semantic as the reasons for enacting moral solutions and political solutions were the same, and political solutions could not be proscribed when they were advantageous.

21. Narveson, *The Libertarian Idea*, 135.

much as possible. But in deciding how much is possible, each is constrained by the recognition that he must neither drive others away from the bargaining table, nor be excluded by them. . . . [And s]ince one wants to benefit from a share of the co-operative surplus, one has no interest in causing the process of bargaining to fail, as its failure would result in no co-operation and no surplus.²²

This problem is essentially a coalition game, where we seek the value of each coalition to each participant. Gauthier's claim that we do not want to drive others away from cooperation assumes that the value to us of a larger coalition including the parties in question is higher than that of a smaller coalition excluding them. His claim that we will limit our claims so as not to drive others away acknowledges the potential that, if our claim is too high, others may find the value to them of a larger coalition that includes us to be lower than that of a smaller coalition excluding us.²³

We can now clarify Narveson's claim that contractarianism includes all those with whom cooperation on mutually acceptable terms is possible. Cooperation with an individual X on terms that all can accept is impossible if there exists a coalition {A} such that X is not a member of {A} and all of the members of {A} find the value to them of forming coalition {A} to be greater than the value of forming any larger coalition that included X. The basic idea is this: we consider what agreement we could reach if we included X, and we compare the utility we expected to derive from that agreement to the utility we expect to derive from agreements that we could form without X. If the expected utility to each of us of forming an agreement without X

22. Gauthier, *Morals by Agreement*, 33–34.

23. For more on coalition games, see Debraj Ray, *A Game-Theoretic Perspective on Coalition Formation* (Oxford: Oxford University Press, 2007).

is higher than the expected utility to each of us from forming an agreement with X, then we are rational to exclude X. In that case, cooperation with X on terms that all could accept is ruled out because we would not accept a less advantageous agreement (including X) when a more advantageous agreement is available (without X).

What does it mean to exclude someone from the social contract concerning moral rights? First, it means that when considering which rules would be mutually advantageous, we need not take that individual's interest into account: if a rule confers an advantage on each of us but a disadvantage on the excluded individual, the rule counts as mutually advantageous to all the relevant parties and becomes a moral rule. Second, it means that we cannot expect the excluded individual to comply with the moral rules we develop without coercion: at least one of the rules will be disadvantageous to the excluded individual, and thus that individual would be irrational to act according to that rule unless offered some external inducements in the form of bribes or punishments.

Thus, there is always a large cost to excluding an individual from the contract. There are four great costs associated with coercing others. The first is detection and punishment. In any situation where a member of the excluded group thinks it possible to violate a disadvantageous proscription without being caught, that individual will do so. Thus, in order for our threat to be effective—that is, to stop the targeted behaviour—we must be capable of detecting a significant proportion of the acts in question, and we must punish perpetrators. Such schemes of detection and punishment, where they are possible, are quite costly. Police forces, courts, and prisons are expensive. They take a great deal of money, time, effort, and resources, and they require some individuals to risk their lives to catch and control criminals. Moreover,

for the sake of detection, we are all required to sacrifice some amount of privacy, and to expose ourselves to the harms associated with corruption and abuse of power in the process. Finally, we are all at risk of being wrongly charged by such a system.²⁴ These costs are quite significant. The more we add proscriptions that are disadvantageous to certain groups to our contracts, the greater the need for these detection and punishment systems, and the greater the costs they bring.

The second cost is prevention. We do not simply want to detect and punish the targeted behaviour; we want to prevent it from happening. Again, this requires detection and preventative force. These costs are quite similar to the costs of detection and punishment more generally, but may require even greater sacrifice of privacy.

The third cost is that of precedent effect. There are two kinds of precedent effect that are relevant. The first I call “tu quoque” precedent: by coercing a specific group, one decides and declares that, in similar circumstances, it is rational for others to do the same. As a member of the majority on the issue in question, one would consider it a cost if coercing on this issue set a precedent that led to the coercion of another minority of which one was a member. For instance, if I was a member of a majority who found homosexuality deeply repulsive because of passages in a religious text that we endorsed, yet I knew that a similar majority was repulsed by my dietary practices because those practices violated passages in the same religious text, I would have to consider that if we attempted to prohibit homosexuality by threat based on the

24. Susan Dimock argues that contractarians must always consider the cost to individual liberty and the risk of punishment associated criminalization when considering enacting criminal laws. (Susan Dimock, “Contractarian Criminal Law Theory and *Mala Prohibita* Offences,” chap. 6 in *Criminalization: The Political Morality of the Criminal Law*, ed. R.A. Duff et al. (Oxford: Oxford University Press, 2014), 177.) Here, I try to specify these costs in further detail. Indeed, while I think that there is conceptual room to discuss exclusion from the contract, with its attendant coercion, from a pre-state moral perspective, the idea of excluding and coercing is very close to that of criminal law, and seems to require that individuals are at least beginning to act collectively in a political manner.

size of our majority, the strength of our distaste, and our religious text, my dietary practices would likely be prohibited for the same reason at some time in the future. Thus, in evaluating whether or not coercing homosexuals was in my interest, I would have to factor in the costs I would bear from having my dietary practices coercively prohibited, and all other similar coercive threats that I would be likely to face. By publicly endorsing the view that coercion under certain circumstances is warranted, I would risk facing coercive threats under all relevantly similar circumstances in which I found myself in the minority. Thus, when considering the costs and benefits of coercion, I should factor in the anticipated costs of the coercion that I would be likely to face. This cost varies: to the extent that the majority is correct in thinking that this instance of coercion is exceptional—that there are almost no other circumstances that are sufficiently similar to call for similar coercion—the majority can consider this cost to be minimal or non-existent.

The second kind of precedent effect is the slippery slope. It is essentially the same as the “tu quoque” except that it takes into account the distortions of human reasoning. The majority may be right that the circumstance is exceptional, yet by engaging in such coercion, other majorities may, due to honest mistake or biased reasoning, mistakenly conclude that they are in a relevantly similar circumstance and thus that it is rational for them to engage in similar coercion. We know that humans are fallible even when they are disinterested, and that they are even more fallible and biased when they are interested parties. As this reiterates, the system snowballs into a Hobbesian state of war of all against all. This is a legitimate slippery slope argument: one is reasonable in predicting such effects and taking them into account as a cost. The greater the anticipated slippery slope precedent effect, the greater the

cost. All of the horrors Hobbes associated with the state of nature—the arresting of science, art, engineering, technology, medicine, etc.—would be significant.

The fourth cost is the foregone benefits of compromise. Were we to compromise with the group in question on principles we find less appealing and include them in the coalition, we could have secured benefits for ourselves. These foregone benefits need to be considered.

Given these four costs, we can expect little coercion: it is generally disadvantageous to try to coerce a group into obeying disadvantageous principles, and thus it will rarely be in our interest to exclude individuals from the social contract. It is generally much more advantageous to compromise. Nevertheless, where the costs identified above are least, a coalition excluding some members may be advantageous to the coalition members. Where the excluded group is very small, weak, and disorganized, and where the targeted action is easy to detect both in advance and after the fact, the costs of detection, prevention, and punishment will be very low. Where the case is truly exceptional, the *tu quoque* precedent effect will be minimal. The slippery slope precedent effect is generally much more dangerous: where society is mature and trust is sufficiently established that this coercion will not be considered a more general threat—the first step in a larger scheme of violence—this cost will be lower. This cost can also be reduced by very carefully and persuasively making the case for this coercion publicly. Moreover, where the targeted minority is not closely tied, either as a subgroup or an ally, to some other minority that will feel threatened, again the cost will be lower. The first three costs are also made smaller by the size of the group in favour of coercion: where the group is nearly all of society, save for the small super-minority being coerced, detection, protection, punishment, and precedent

effect costs are lowest. Finally, where the benefits of compromise are least, the case for coercion will be stronger.

When all of these costs are at their lowest, it seems reasonable to think that excluding a group will be in the interest of coalition members. In such cases, a proscription that is disadvantageous to a small minority might be added to the contract, compliance with which is to be *made* advantageous for the minority by threat of punishment. We can now replace Hobbes's and Hume's claims that people do not want to harm others non-instrumentally with the claim that it will generally be rational to exclude from cooperation people who want to harm others non-instrumentally. By thus excluding such individuals, we are able to secure moral rights like the right to life: insofar as the relevant parties to the contract want to not be killed, and do not want to kill others non-instrumentally, a rule proscribing killing others will be mutually advantageous.

3.1.3 Why Consider a State of Nature?

As we saw in section 2.1.1, Hobbes's reason for examining a state of nature was based on his endorsement of the resolute-compositive method, according to which the best way to explain a phenomenon is to examine individually the nature of each of the parts contributing to that phenomenon, and then to consider how those parts come together to produce the phenomenon.²⁵ This commitment is unlikely to strike contemporary readers as persuasive; moreover, contemporary philosophers have argued that moral and political theories should disregard the state of nature, focusing instead on other theoretically-specified non-agreement points.²⁶ A new argument for examining a state

25. Page 19.

26. See, for instance, Barry, *A Treatise on Social Justice, Volume 1: Theories of Justice*, 370.

of nature is needed.

Such an argument is provided by Jean Hampton in a political context that can be reformulated for moral arguments. Hampton claims that historical and hypothetical interpretations of social contract theory are partially right and partially wrong. The hypothetical interpretation is wrong insofar as contract theory seeks to explain the *actual* relation between citizens and sovereigns; the historical interpretations are wrong insofar as contract theory seeks to *justify* the existence of a certain form of state and historical promises are not up to the task.²⁷ Rather, the contract theory should be interpreted as “a certain kind of *rational reconstruction* of the state *with historical import*.”²⁸ The contractarian story of how and why a state emerges from a state of nature applies not just to the *creation* but also to the *maintenance* of the state: state power is created and maintained by the obedience of the citizens—by their consent—whether or not the citizens or the sovereign believe this to be the case.²⁹ Though we may obey the king because we believe him to be divinely chosen, it is nevertheless our obedience, and not some divine preference, that gives the king his power.³⁰ We can revoke this consent at any time by refusing to obey, returning to the state of nature. Moreover, because contract theory seeks to *justify* the existence of a state of a certain form, it appeals not to the actual historical beliefs that led to the creation and maintenance of the state—for instance, the belief that the sovereign is divinely chosen—but rather to the rational interests of the citizens.³¹ Thus, while there are idealized aspects to contract theory, like the appeal to rational interest instead of actual beliefs, the contract nevertheless aims to capture the relation between

27. Hampton, *Hobbes and the Social Contract Tradition*, 266–268.

28. *ibid.*, 269.

29. *ibid.*, 272–273.

30. *ibid.*, 273.

31. *ibid.*

citizens and sovereigns in reality, not hypothetically.

This political interpretation of contract theory can be revised to apply to the contract theory of moral rights discussed in this thesis. Moral rules have their force because we obey them; they are conventional. Should we refuse to obey them, we can bring about a non-cooperative state of nature. The contract theory justifies certain moral rules by showing that it is rational to follow them, disregarding the beliefs that actually bring people to obey purportedly moral rules, like the belief that the rules are divine commands, or that some rules are simply obviously true. Such arguments are important precisely because it is in our power to bring about a non-cooperative state of nature, which is bad for everyone, and because there are specious but commonly endorsed patterns of reasoning that can lead people to promote such a state of nature. The contract argument explaining why in a state of nature we should create a moral convention justifies both the creation and the maintenance of the specified moral rules; thus, the argument is instructive even if we do not presently find ourselves in a state of nature.

3.1.4 The Revised Contractarian Theory of Moral Rights

We can now re-articulate the Hobbesian mutual advantage contract theory of moral rights in light of the revisions presented in sections 3.1.1 through 3.1.3. The theory starts by justifying the creation of a moral convention in a non-cooperative state of nature because moral obligations truly are conventional—they depend on our willing obedience, and are not natural—and because such an argument specifies the moral rules that we ought to obey and justifies our continued obedience. This is why Hobbes insists that all constraint must be justified against the baseline of our unlimited right

of nature.³²

The theory seeks to *justify* moral constraints to as many people as possible, subject to the constraint examined in section 3.1.2. Thus, it starts with as broad an account of human nature and rationality as possible. On this account, each individual wants that which is most pleasing and least displeasing, or at least wants as many of his or her most important desires to be satisfied as possible. Humans generally prefer to continue to live, and almost always prefer not to be killed against their will. Human desires are variable, both from time to time and from person to person. Humans generally do not want to harm other humans non-instrumentally, and those few who do ought to be excluded from the contract.³³

Against the baseline of a non-cooperative state of nature, the theory presents two sources of constraint: the laws of nature and obligation.³⁴ Hobbes justifies the laws of nature by showing that, given human nature, each individual must find these laws to be in their interest. We obey the laws because obeying the laws will best preserve our lives and promote our preference satisfaction, and as section 3.1.2 argued, we strongly desire not to be killed against our will. These laws of nature are natural insofar obeying the laws of nature when the circumstances of justice hold will always be the best way to preserve our lives and promote interests.³⁵ We should obey these laws because doing so is the best way to promote the interest we take in our continued lives and our happiness. If we did not desire the effects to which they led us, then they would provide us with no normative principles of action.

32. This argument summarizes Hampton's argument, modified in section 3.1.3 to apply to moral theories. Hampton, *Hobbes and the Social Contract Tradition*, 266–273. For more on convention, see section 2.2.2.

33. See section 3.1.2.

34. Discussed in sections 2.1.5 and 2.1.6.

35. For more on the circumstances of justice, see section 2.2.1.

The laws of nature direct us to form a convention for the sake of peace, thereby obliging ourselves (I.xiv.5). The contract theory specifies the articles of the convention by reference to mutual advantage: all those moral principles that are mutually advantageous to all of the relevant parties should be part of the convention. The relevant parties are determined by a coalition game which leads us to exclude all those with whom cooperation would be disadvantageous for us, for instance, people with misanthropic desires.³⁶

Such a contract theory justifies all constraints by showing them to be in our own interest. The laws of nature are natural causal knowledge, but they provide us with reasons for action only insofar as they show us the most efficient means to the things we actually desire. And we carry out our obligations only because carrying them out is in our interest: were it ever in our interest to violate a contract, we could and should do so. However, as the third law argues, such a situation could not arise—practically speaking, it will always actually be in our interest to keep our contracts. Thus all constraints on our behavior are derived from our prudential interests.

While such a Hobbesian argument offers justifications for constraints that are relative to each individual, it is able to generate one set of constraints for all people because it focuses on what is common to all the humans with whom cooperation is possible.³⁷

Thus, the contract theory justifies all constraints, whether they derive from the laws of nature or from convention, by showing them to be mutually advantageous in the circumstances of justice. But as I mentioned in section 2.2.1, the circumstances

36. This argument was developed in section 3.1.2.

37. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*, 90.

of justice—in particular that of the rough equality of the parties to the convention—can hold to a greater or lesser extent, and it is unclear when the circumstances hold too weakly for moral principles to obtain. The rest of this chapter focuses on rough equality, and asks how weakly that constraint can hold and yet still give rise to moral constraints. It examines the precise abilities required to qualify directly for moral rights, as well as the conditions that would give rise to indirect moral protections.

3.1.5 Mutual Advantage Contractarianism and Rights

What does it take to qualify directly for rights according to mutual advantage contractarianism? Before answering this question, I should like to say that my answer is not novel, except perhaps in its detail, and is the same as Narveson's, who is a prominent mutual advantage contractarian and a fierce critic of animal rights. I do not take myself to be disagreeing with Narveson on any point about the logic of contractarianism: I think that the contractarian logic that Narveson endorses is largely correct and ought to have led him to endorse some animal rights.³⁸ Moreover, while my arguments were devised independently, I have since found similar arguments that either lacked the same detail as mine, or differed in details, by C. Tucker & C. MacDonald, Christopher W. Morris, and Andrew I. Cohen.³⁹

38. Narveson, "Animal Rights," 177, Narveson, "Animal Rights Revisited," 57, Narveson, "On a Case for Animal Rights," 42.

39. Tucker and MacDonald, "Beastly Contractarianism? A Contractarian Analysis of the Possibility of Animal Rights"; Morris, "Moral standing and rational-choice contractarianism"; Morris, "The Idea of Moral Standing"; Cohen, "Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals"; Cohen, "Contractarianism and Interspecies Welfare Conflicts." Where we agree, I take this agreement as reassuring, and develop the ideas in greater detail. Indeed, I think my argument is largely in agreement with Tucker and MacDonald, differing only in that I add greater detail (which I think brings greater clarity) and that I go on to defend the contractarian position. Where we disagree, I will defend my views, particularly in section 3.1.7.

Returning to the question, I claim that, because rights and duties, on the contractarian view, are grounded in rational mutual advantage, the question of what it takes to have rights must be answered from the perspective of the agent whose behaviour is to be restrained by the rights in question: with respect to whom would it be in the agent's interest to restrain his or her behaviour? Thus, to qualify for moral rights on the mutual advantage contractarian view, (a) one must be capable of responding appropriately to the behaviour of others,⁴⁰ (b) one must be capable of imposing costs *in response* to their unrestrained behaviour,⁴¹ and (c) those costs must outweigh the benefits to others of their unrestrained behaviour.⁴² These criteria are individually necessary and jointly sufficient to guarantee direct moral rights, on the contractarian view.⁴³

In order to qualify directly for a right, (a) one must be capable of imposing some cost on those who would violate one's purported rights, and (b) one must do so in response to that violation. The relevant sense of "cost" here is that of "opportunity cost," which includes not only the ability to impose some harm on the actor, but also the ability to withhold some benefit from the actor. For example, a small cat may be incapable of causing a strong human adult serious harm, but may be able to refuse to allow the human to pet it and to avoid being near the human: the cat's affection

40. Tucker and MacDonald call this requirement the "responsiveness" requirement. (Tucker and MacDonald, "Beastly Contractarianism? A Contractarian Analysis of the Possibility of Animal Rights," §V.1)

41. Tucker and MacDonald label this requirement the "potency" requirement. (ibid.)

42. Tucker and MacDonald seem to overlook this consideration of what might be termed *relative* potency. As Morris notes, the properties grounding moral status for contractarianism are relational. (Morris, "Moral standing and rational-choice contractarianism," 77)

43. I think that this way of explaining the capacities required to qualify directly for moral rights is clearer than, though largely in agreement with, the criteria that Morris offers, according to which parties hold rights against one another if and only if (a) they are in the circumstances of justice with respect to one another, (b) they are capable of restraining their behaviour, and (c) they are willing to restrain their behaviour. (ibid., 86–87)

would be a foregone benefit. If the human is a cat-lover who craves some sort of affectionate interactions with the cat, then the cat's ability to withdraw itself from voluntary interaction could impose a cost on the human. If, however, the individual in question lacked the ability to impose any cost on the agent in question, and there was any positive value at all to be gained from acting without restraint toward that individual, then such an individual could not qualify for rights. For in such a case, it could never be to the advantage of the agent to act with restraint toward the individual: there would be advantage to be had from acting unrestrainedly and no comparable disadvantage to such behaviour. Thus, the ability to impose a cost on the purported duty-bearer is requisite for an individual to qualify as a right-holder.

However, the ability to impose cost, on its own, is insufficient to qualify one as a right-holder. It is also necessary that one be capable of modifying one's behaviour in response to the actions of others. One must be capable of imposing a cost in response to unrestrained behaviour, and refraining from imposing a cost in response to restrained behaviour. For instance, consider mosquitoes: mosquitoes have the ability to be severe annoyances to humans. I, for one, would be happy to refrain from killing mosquitoes, and to supply them with a constant stream of alternate sources of nourishment, if they would all stop biting me. But the mosquitoes cannot refrain from harming in exchange for restraint. Thus, I have no reason to observe restraint toward the mosquitoes. If the individual in question lacks the ability to modify its behaviour in response to mine, then any cost that it is capable of imposing on me is as likely to follow from my restrained behaviour as it is to follow from my unrestrained behaviour, and this ability to impose cost would fail to count in favour of my restrained behaviour.

Yet it may be the case that an individual is capable of responding to the unrestrained behaviour of another by imposing some cost on that other, and yet that the individual fails to qualify for rights. A further requirement is that the cost that the individual can impose on the other outweigh the benefits to the other of her unrestrained behaviour. Thus, ladybugs of the appropriate variety may often decide to bite my young nephew for his curious handling of them, yet that bite is never sufficient to dissuade him from picking them up and playing with them. This requires a certain contingent relationship between the purported rights-bearer and the purported duty-bearer that is subject to change: perhaps bees' stings were sufficient to dissuade humans from invading their hives for honey at one point, but technological advances have allowed humans to avoid the bees' stings, divesting the bees of any rights they could have had. Once humans find themselves without the relevant technology, they revert to respecting bees' hives.

Because of the way that this argument relies on the abilities of the purported right-holder relative to those of the purported duty-holder, rights justified in this way do not yet entail third-party obligations. This account of rights explains how the duty-bearer acquires duties toward the right-holder, but does not justify obligations for governments or other third parties to help enforce the rights of the right-holder, or to punish violations of those rights. A further argument must be given, showing when it is in the interests of third parties to endorse enforcement and punishment schemes. This argument can often be given, but may not apply to all right-holders.

To sum up the discussion up to this point, in order to qualify directly for moral

rights on the mutual advantage contractarian view, (a) one must be capable of responding appropriately to the behaviour of others, (b) one must be capable of imposing costs in response to their unrestrained behaviour, and (c) those costs must outweigh the benefits to others of their unrestrained behaviour. These criteria are individually necessary and jointly sufficient to guarantee direct moral rights, on the contractarian view.

3.1.6 Rights v. Protectorate Status

Up to this point I have been focusing on the criteria for direct moral rights, and I have been rather vague about what ‘qualification for direct moral rights’ entailed. I would now like to clarify what I mean, and to contrast direct qualification for moral rights with what I will call ‘protectorate status.’ To qualify directly for rights is to pose sufficient threat to others that it is in their interest to take your interests into account when trying to find mutually advantageous principles. Such principles need to be to the advantage of all individuals who qualify directly for rights; whatever principles of restraint such individuals find mutually advantageous will determine their rights. To qualify for protectorate status does not entitle one to have one’s interest considered directly; rather, protectorate status arises when those who qualify directly for rights decide that it is to their mutual advantage to impose restraint on their behaviour toward other groups who do not qualify directly for rights.

Another way to explain the distinction between direct moral rights and protectorate status is to draw attention to the individual to whom the duty-bearers owe their duties. On the one hand, individual J has a direct right against individual K for X to happen if and only if K has a correlative duty owed to J not to interfere with

the occurrence of X. On the other hand, individual L has a protection guaranteeing the happening of X, stemming from L's protectorate status, if and only if individual M has a duty to individual N to not interfere with the happening of X. In the case of direct rights, the duty holder owes a duty directly to the right-holder, and that duty is owed based on the abilities of the right-holder relative to the abilities of the duty-holder, as explained above. In the case of protectorate status, the duty not to interfere with the protected individual is owed to some other individual who has the requisite relative abilities. In this case, the protectorate status is parasitic on the direct right of the other individual: should the direct right-holder perish or lose the requisite relative abilities, the protectorate status would disappear.

3.1.7 Protectorate Status

'Protectorate status' arises when one individual has certain interests but does not qualify directly for rights, and at least one other individual who does qualify for rights decides to advance claims to secure the protected individual's interests. Thus, the set of protectorates could be viewed as a subset of the things that a right-holder would endeavour to 'protect': the subset of those things who take an interest in their own well-being. A right-holder would attempt to defend his or her inanimate property, imposing duties on others not to damage or otherwise interfere with that property, but inanimate property cannot take an interest in its own well-being and such protection cannot be said to be in the interest of the property. By contrast, a protectorate is a living individual that takes an interest in its own well-being, and whose interests the protector endeavours to protect.

In section 1.1 I established a preliminary definition of "moral right" that involved

two criteria, the second of which required that the right-holder benefit from the right. By claiming that protectorate status applies only to individuals who take an interest in their own well-being I am incorporating the second criterion from the definition of a right into the definition of protectorate status. Thus, not only must a protectorate take an interest in its own well-being, but all protections must benefit the protectorate in question. By defining protections to be in the interest of protectorates, I do not intend thereby to prohibit actions that are against the interests of the protectorates. Rather, in order to count as a “protection,” the rule in question must benefit the protectorate; however, that does not entail any proscription of rules concerning protectorates that are not to their benefit.

The concept of protectorate status is thus close to being encompassed in the concept of property, but there are certain important and subtle differences between protectorate and property status other than the difference of taking an interest in one’s own well-being. Firstly, when one claims property rights to something, one generally claims the exclusive rights to do a host of things to or with that property. When I claim a property right to my car, I claim to be the only one who can destroy that car, the only one who can drive that car, the only one who can sell that car, etc. Thus a property right to X is generally a bundle of exclusive rights to do things with X. Moreover, claims to a property right are generally grounded in a historical transaction: I have a property right to X because I acquired X legitimately. Thus a property right is only indirectly grounded in mutual advantage: the rules for acquiring and exchanging property, and for respecting the property of others, will be grounded in mutual advantage, and it is through those rules that one asserts a right to a particular piece of property. By contrast, a claim to protectorate status for X need

not be grounded in any historical transaction: I can claim that my infant, or any other infant for that matter, has protectorate status without needing to claim that I acquired that infant in any way. My claim of protectorate status will be grounded directly in mutual advantage: I claim that I will punish violations of the interests of the infants so thoroughly that whatever benefit others would receive from those violations would be outweighed by the consequences of the punishment. Moreover, when I claim protectorate status for X, I do not claim the host of exclusive rights typically associated with property rights. Firstly, I do not claim the exclusive right to violate the interests of X; I claim that no one should violate the interests of X. Secondly, by advancing a protectorate claim for X I need not claim the exclusive rights to control the whereabouts of X, or to play with X, etc. I may also choose to advance a property right to X, perhaps feeling that this is the best way to guarantee the protection of X's interests, but I need not do so. Thus, if I claimed protectorate status for some remaining herd of an endangered species, I would not be asserting any exclusive rights to treat that herd in any given way, such that anyone wanting to study, observe, or interact with the herd in a beneficial manner would need my authorization. I would simply be claiming that others are forbidden from violating the interests of the herd.

This "protectorate status" may have many of the same drawbacks as property status, and may be open to many of the same objections; however, the concept is subtly and importantly different from property status, and I hope to have exposed those differences. If this way of talking about protectorate status sounds too strange, then one can translate protectorate status claims to claims about the indirect duties of others: "individual X is protected in the doing of Y" translates to "others have

a duty to some individual (or set of individuals) Z not to disrupt X's doing of Y." Perhaps my claims will sound less peculiar presented in this manner.

To have protectorate status does not require any ability to respond appropriately to the behaviour of others, or to impose sufficient costs on others. Because protectorate status is parasitic on actual rights, one must examine the direct right-holders and the kinds of principles that they would find mutually advantageous to determine who would qualify for protectorate status, and what protections would follow from that status. Right-holders may find it mutually advantageous to refrain from cruelty toward animals, for instance, but may find that the drawbacks from refraining from eating animals are too considerable to be mutually advantageous. In order for some J to have a protectorate status, it would have to be that the agents found it rational—viz., the means to best fulfill their desires—to treat J with the appropriate restraint. This requires that there be some individual (or set of individuals) K with the abilities requisite for direct rights who cares enough that J be treated with restraint to threaten retaliation against any unrestrained behaviour. So the indirect duty-holders must find that the benefits of unrestrained treatment of J are outweighed by the costs K would impose combined with any potential benefits that the duty-holders could reap from a convention protecting individuals like J from unrestrained treatment. And K must find that the costs of threatening retaliation, and of carrying out that retaliation, are outweighed by the benefits of having J treated with restraint. The more that the purported indirect duty holders found that they would stand to gain from a convention recognizing such protectorate status, the lower the costs that K would need to be capable and willing to impose in retaliation for unrestrained behaviour, and the less inclined K would need to be to do so. Should all

duty-holders find that they stand to benefit from recognizing the protectorate status in question—that the benefits they reaped from such restraint on each duty-holder's part outweighed any gains they stood to reap from unrestrained behaviour, whether or not anyone retaliated—then there would be no need for anyone to be capable and willing to carry out threats of retaliation against unrestrained behaviour for such protectorate status to come about.

The paradigmatic case of a protectorate status requiring no threat is that protecting humans with disabilities. All right-holders may decide that it is in their interest to observe restraint with respect to humans with disabilities insofar as each right-holder faces a very high probability of being disabled for some portion of his or her life, and thus a very high probability of benefiting significantly from such restraint in the future. On the other hand, the paradigmatic case of a protectorate status that requires threat is that protecting children too young to qualify for rights directly. Parents strongly desire that their young defenseless children not be subject to any abuse or attack. If they were strongly motivated to retaliate by imposing severe costs against anyone who should fail to act with restraint toward their children, and if the costs they would impose outweighed the potential benefits that the purported indirect duty-holders would reap from unrestrained treatment of the children in question combined with the potential benefits that the purported indirect duty holders would reap from the general recognition of such protectorate status, then the protectorate

status of these children would be established.⁴⁴

Protectorate status is a subset of the kind of coalition games discussed in section 3.1.2.⁴⁵ A super-majority will form a coalition and exclude or threaten a minority whenever the exclusion or threat is expected to yield greater benefits than harms to the majority's members. Protectorate coalition games are those coalition games where the super-majority in question derives its relevant expected utility from the way some other group of sentient beings is treated. Such a threat will only be made where the benefits outweigh the costs, but that should rarely happen. Thus, in general, few protectorate threats would be likely to be undertaken. Nevertheless, some will. The protections afforded to children fall under the category of protectorate threat.

Protectorate status is similar to what has been called "secondary moral standing," but differs in a subtle way.⁴⁶ An individual has moral standing if and only if it is directly owed moral duties.⁴⁷ Primary moral standing entails qualifying for rights on

44. This argument for parents enforcing protections for their children is discussed Chapter 15, "Children and Rights" in Narveson, *Respecting Persons in Theory and Practice: Essays on Moral and Political Philosophy*, 265–278; and in Chapter 19, "The Problem of Children" in Narveson, *The Libertarian Idea*, 269–274. Such protections would not, however, defend children against their own parents. Children would be protected against their parents by protections established by the broader community. Because children will grow to potentially threaten human lives, community members have an interest that children should be brought up to cooperate peacefully and, perhaps, to be contributing members of the community. Any actions on the part of parents that threatened this interest—for instance, by abusing, neglecting, manipulating, or deluding their children—would be proscribed. This argument is developed in Chapter 15 of Narveson, *Respecting Persons in Theory and Practice: Essays on Moral and Political Philosophy*, 265–278. This line of reasoning will be deployed in section 3.2.2 to argue that dogs should not be abused or starved.

45. Page 82.

46. Morris, "Moral standing and rational-choice contractarianism," 76; Morris, "The Idea of Moral Standing," 266–267; Cohen, "Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals," 188; Tucker and MacDonald, "Beastly Contractarianism? A Contractarian Analysis of the Possibility of Animal Rights," §IV.

47. Morris, "Moral standing and rational-choice contractarianism," 76, 82; Cohen, "Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals," 188; Cohen, "Contractarianism and Interspecies Welfare Conflicts," 234–235.

one's own merits in the manner detailed in section 3.1.5.⁴⁸ Secondary moral standing entails deriving one's moral status from others who qualify directly for rights based on their interests: a group of right-holders—call it *A*—refuses to cooperate unless secondary moral standing is afforded to group of non-right-holders *B*.⁴⁹ Protectorate status and secondary moral standing differ in two regards.

The first difference between protectorate status and secondary moral standing lies in this detail: when *A* refuses to cooperate unless *protectorate status* is extended to *B*, they are only holding out for indirect moral duties and indirect moral status for *B*. Right-holders would have duties owed to *A* to respect certain interests of *B*. Whereas, when *A* refuses to cooperate unless *secondary moral standing* is extended to *B*, they are holding out for direct moral duties and direct moral status for *B*. Right-holders would have duties to *B* to respect certain interests of *B*.⁵⁰ I think that this difference is largely academic: there is no practical difference between secondary moral standing and protectorate status—the difference is purely theoretical. Both have the same “defeating” conditions: should *A* stop taking an interest in the wellbeing of *B*, or should all of the members of *A* die, *B* would lose both secondary moral standing and protectorate status.⁵¹ Given that the strength and content of the duties are equal,⁵² and the defeating conditions are identical, it seems that secondary moral standing

48. Morris, “Moral standing and rational-choice contractarianism,” 89–90; Cohen, “Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals,” 190; Cohen, “Contractarianism and Interspecies Welfare Conflicts,” 235.

49. Morris, “Moral standing and rational-choice contractarianism,” 89–90; Cohen, “Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals,” 190; Cohen, “Contractarianism and Interspecies Welfare Conflicts,” 236.

50. Morris, “Moral standing and rational-choice contractarianism,” 89–90; Cohen, “Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals,” 190; Cohen, “Contractarianism and Interspecies Welfare Conflicts,” 235–236.

51. Cohen, “Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals,” 193–194.

52. “Content,” here, must obviously exclude the question of whether the duty is *to B* or merely *concerning B*.

and protectorate status are practically the same. Given this practical identity, I contend that we should be suspicious of secondary moral standing. There seems to be an assumption that protectorate status is somehow weaker or less interesting than secondary moral standing.⁵³ But consider if *A* was negotiating consideration for *B* with other right-holders, and the other right-holders said “we will grant *B* protectorate status, but not secondary moral standing.” It would make no sense for *A* to refuse that offer and hold out for secondary moral standing: the interests of *B* are just as well protected by protectorate status as by secondary moral standing. Thus, I think that the idea of secondary moral standing is a theoretical distraction. It simply makes more sense, given that a relevant change in *A*’s interests or the elimination of *A* would annul all obligations to *B*, to say that such consideration is indirect and thus amounts to protectorate status.⁵⁴ To insist on describing things otherwise is to insist on describing things in a misleading way for no relevant practical reason.

The second difference is that some philosophers explicitly take secondary moral standing to entail that those with secondary moral standing are owed *all* of the same obligations that are owed to those with primary moral standing.⁵⁵ Insofar as secondary moral standing must guarantee protections equal to those of primary moral

53. “An intermediate case would be one where Frederica would be given the status of an indirect moral object. But such a case is not theoretically interesting.” Morris, “Moral standing and rational-choice contractarianism,” 89.

54. *Pace* Cohen, “Contractarianism, Other-regarding Attitudes, and the Moral Standing of Non-human Animals,” 193–194.

55. Tucker and MacDonald explicitly endorse this claim. (Tucker and MacDonald, “Beastly Contractarianism? A Contractarian Analysis of the Possibility of Animal Rights,” §IV) It is less clear whether Morris endorses this view; he considers an objection from James Rachels to the effect that such singular notions of “moral standing” are unhelpful and seems to think that it applies to his view, but does not find it detrimental. (Morris, “The Idea of Moral Standing,” 270). Likewise, Cohen is not explicit on this matter. (Cohen, “Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals”) Thus, my discussion of this difference between secondary moral standing and protectorate status should be taken to apply only to those theories that actually endorse this interpretation of secondary moral standing.

standing, far fewer individuals will qualify for secondary moral standing. In that case, we would need protectorate status to account for our obligations toward those who no longer qualified for secondary moral standing. Because I do not see a practical reason to prefer secondary moral standing to protectorate status, and because protectorate status is more inclusive than secondary moral standing (insofar as secondary moral standing demands protections equal to those of primary moral status), I will focus on rights and protectorate status, disregarding secondary moral status.

3.1.8 Summary

To sum up this section, an individual qualifies directly for rights if and only if (a) the individual has the ability to respond appropriately to the actions of others, (b) the individual has the ability to impose opportunity costs on others, and (c) the opportunity costs that the individual can impose on others outweigh the benefits those others derive from acting without restraint toward the individual in question. Alternately, an individual qualifies for protectorate status if and only if either both (a) someone else qualifies directly for rights, and (b) that person would be happier threatening and perhaps carrying out threats of retaliation against others who would act without restraint toward the individual in question than that person would be allowing others to act without restraint toward that individual, or else (c) all purported duty-bearers find that the benefits of general restraint toward such individuals outweigh the benefits of unrestrained behaviour toward such individuals. However, protectorate status will only arise when the attending costs of detection, punishment, prevention, and precedent effect are minimal, and where the forgone benefits of compromise are negligible.⁵⁶ Thus, for one to qualify directly for rights is for one to pose sufficient threat

⁵⁶. These costs are examined in greater detail in section 3.1.2.

to others to make it in their interest to take one's interests into account when they are deciding how to act. For one to qualify for protectorate status, those who qualify directly for rights must decide that it is in their interest to act with restraint toward the individual in question.

3.2 Direct Moral Rights for Animals

Given the criteria above, do any animals qualify directly for moral rights? In order to determine whether any animals qualify directly for rights, we need to know (a) whether the animals respond appropriately to the behaviour of others, (b) whether they can impose costs on agents, and (c) whether the costs they can impose outweigh the benefits that agents derive from unrestrained behaviour. The 'agents' in question are those individuals who qualify directly for rights. From the outset we know that able humans will qualify directly for moral rights, and it is human behaviour that we are trying to modify by asserting animal rights, so we must identify the benefits and harms that animals can give to, or withhold from, able humans.

Even with the required abilities laid clearly before us, it is difficult to determine in more than a suggestive way which animals are entitled to rights because of the massive amounts of information required to make such a determination: one must take into account, and balance against one another, the prevalence and strength in humans of various desires involving animals, the various desires that a given animal is capable of fulfilling, the levels of harm that animals can impose on humans, and the abilities of animals to withhold those benefits in retaliation for ill treatment and to withhold those harms in response to restrained treatment. And insofar as different individual members of given species will take an interest in different things, we need to

know about the desires of individual animals.⁵⁷ Thus a vast knowledge of human and animal natures is required to decide precisely which animals are entitled to precisely which rights, and individual variability entails that this examination is best carried out individual by individual. To offer general guidelines for animal rights, we must attempt to categorize individual animals in ways that focus on their possession of the abilities relevant for rights.

There are two different ways to categorize animals to this end, neither of which is completely satisfactory, but both of which, when combined, offer a more complete picture of what animal rights are called for by mutual advantage contractarianism. The first way to categorize animals is by the benefits that humans derive from the animals, and the second is by species. Neither is fully satisfactory because the categories conceal great variation in the relevant abilities among individual category members: while a German shepherd might have all of the abilities relevant for rights, a chihuahua, though equally a member of the species ‘dog,’ and a hamster, though equally a member of the category ‘companionship animals,’ may not. By combining these two categorizations, we can offer more accurate guidance on the rights owed to a certain individual animal, though this guidance will always need to be supplemented with further information about the individual in question and attention to the principles detailed in section 3.1.5. Section 3.2.1 will begin by categorizing animals according to the benefits that humans derive from them, and examining the rights entailed by each category. Section 3.2.2 will then examine some animal species and the rights

57. Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 111, discussing Elizabeth Marshall Thomas, *The Hidden Life of Dogs* (Boston: Houghton Mifflin, 1993).

owed to them.⁵⁸

3.2.1 Benefits

There are at least nine different benefits that humans derive from animals: (1) companionship, (2) security, (3) hunting assistance, (4) transportation, (5) entertainment/aesthetics, (6) medical/testing, (7) nourishment, (8) clothing and other products, and (9) hunting.⁵⁹ Some of these categories may overlap slightly, and many animals are capable of providing several of the benefits listed. As we will see, these benefits are listed on a spectrum from those benefits that count most strongly in favour of the broadest set of rights to those benefits that count most strongly against rights. In the middle are benefits that count for basic rights, and may count for or against further rights depending on the specifics of the case. I will examine all of these benefits below, but due to structural similarities, hunting assistance, transportation, and entertainment will be addressed together under the heading “service.”

Before examining each of these benefits, it must be noted that not all humans may desire such purported benefits from animals—some humans may not want animal companionship, and other humans may prefer to starve than to eat animals. Thus, each of these purported benefits will count for or against rights *only in proportion* to the strength and prevalence of the desire for that benefit among humans. The greater the number of humans who want companionship from animals, and the stronger

58. One could also categorize animals according to the kinds of harms they can impose on humans; however, such a categorization offers much less useful guidance insofar as, for the most part, the animals capable of imposing severe harms on humans are incapable of withholding those harms in response to appropriate treatment. Most of the relevant costs that animals can impose on humans come from withheld benefits, not from positive harms. For that reason, I forgo categorization based on harms.

59. Donaldson and Kymlicka identify a number of the items in this list. See Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 73.

that desire relative to their desires for whatever benefits they could obtain through unrestrained behaviour, the stronger the case for rights for animals capable of offering companionship. Inversely, the greater the number of humans who want to eat animals, and the stronger their desire to do so, the stronger the case against rights for tasty animals. This qualification is assumed to apply to all of the benefits discussed below. Moreover, animals belong to these categories only so long as they continue to provide the benefit in question. Should an animal cease to provide the relevant benefit, or should the human no longer desire that benefit from the animal in question, the rights associated with the provision of that benefit would no longer be owed to the animal in question.

Companionship

Many humans seek out the companionship of domesticated animals, desiring their affection. But we do not just want the animal to *feel* affection for us—we want that affection to be communicated, which is accomplished by certain behaviours: purring, licking, allowing to be petted, playing, etc. How many humans would identify companionship as a benefit that they derive from animals? This is up for debate. Nussbaum feels that the ability to share the companionship of animals is a basic component of human dignity and flourishing, and that a society that fails to guarantee the capability for humans to establish these relations with animals fails to be minimally just.⁶⁰ Yet perhaps some people do not feel any strong desire to relate with animals in this way.⁶¹ Again, the more wide-spread this desire is, and the stronger it

60. Martha C. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership* (Cambridge, Mass.: The Belknap Press: Harvard University Press, 2006), 75–77.

61. See Fran Liebowitz's view, discussed in Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 211–212.

is, the more considerable are the benefits that companionship animals can offer, and the stronger the foundation it offers for direct rights claims by those animals if they can withhold that benefit in retaliation for ill treatment.

It would seem that many companionship animals are capable of withholding their manifestations of affection in response to what they consider to be ill treatment. Most pet owners can recount some time when their pet took exception to something the owner did, and proceeded to withdraw from the owner, or engage in some other form of retaliation.⁶² Whenever my parents travel and leave our dog in someone's care, upon my parents' return, our dog will avoid my parents in the house as much as possible for a time. It is unlikely that all animals who serve as companions can withhold their affection in this manner, but the more capable they are of withholding their affections, the stronger is their claim to rights.

Insofar as (a) humans value the manifestation of affection from companion animals, (b) those animals are capable of withholding that value to retaliate against ill treatment, and (c) the withheld benefits outweigh the benefits of the ill treatment, those animals will qualify directly for some rights. These rights vary from the basic rights to adequate nutrition, housing, and health care, to more complex rights that must be negotiated between the animal in question and its owner(s). Should a companionship animal care greatly about something—so much so that it will withdraw its affection unless it gets what it wants—and should the owners find it more in their interest to comply with the animal's wishes than to resist, then that animal will have a right to the object in question. While the particular set of rights varies based

62. For examples, see Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 108–112, 115, 117–118, 120. Indeed, Donaldson and Kymlicka point out that domestication presupposes and reinforces animals' abilities to have and communicate a subjective good, to cooperate, and to "participate in the co-authoring of laws." (p. 103–105)

on individual interests and abilities, companionship will count in favour of animal rights whenever an animal is capable of withholding its affections in response to ill treatment.

Security

Some humans seek out animals for the sake of security. People who live alone, or who live in the country, may find it beneficial to keep the company of large dogs who will alert the person to any intruders, and who may indeed frighten or attack any intruders. Moreover, some people simply feel an increased *sense* of security at having a friendly animal present, whether or not that animal would be of any practical use in the presence of some intruder, though such a sense of security fits better with companionship than the kind of security discussed here.

There would seem to be at least two ways to persuade an animal to provide one with security: by earning its loyalty through kindness, or by so impressing it with one's superior physical and mental strength, through beatings or other punishments, that the animal obeys one's commands out of fear. These strategies may not be equally effective: security earned through punishment will last only as long as one's superiority is maintained. Should one fall ill, one might face a revolt on the part of the security animal. Moreover, there must be limits to the extent to which security is purchased through cruelty: unless the animal is provided with the conditions necessary for it to thrive physically, it will not maintain the strength and health necessary to provide the desired security.

Thus, regardless of the relative effectiveness of the two strategies, animals used for security should have basic rights to sufficient nutrition, housing, and health care

for the animals to provide the desired security. Should the latter strategy of cruelty prove more effective than the former strategy of earning an animal's loyalty through kindness, security would count against further rights. Should the two strategies prove equal, security would neither count for nor against further rights: the benefit could be obtained through kindness or through ill-treatment. Should security be better purchased through kindness than cruelty, it would count in favour of further animal rights: security animals would then provide a benefit only in response to favourable treatment. One would need to negotiate with those animals for their continued service, providing whatever benefits they demanded in exchange for providing security, so long as the costs of providing those benefits were outweighed by the benefits.

Service

Under this heading I will address hunting assistance, transportation, and entertainment. Some humans use animals to hunt: my brother was recently looking for a low-budget apartment and was told by a potential land-lord that she preferred her tenants to have cats to keep the mice away. More typical examples of animals helping humans to hunt might be fox-hounds and other scent hounds, retrievers, and pointers. Likewise, humans benefit from predators who hunt pests or other undesirable insects or animals. Sometimes these predators are native to the human environment in question, and other times the predators are deliberately introduced by humans to control other populations.

Animals can also be used for transportation, from horse-drawn buggies to dog sleds. Finally, some animals can be used for entertainment, such as animals in zoos and circuses. There are also other aesthetic benefits that humans derive from animals:

people generally find at least some of the animals living around them aesthetically pleasing: squirrels, birds, raccoons, fox, etc. But these latter aesthetic values do not count in favour of animal rights insofar as the animals cannot withhold these benefits, and are distinct from other forms of animal service insofar as they do not depend on any particular behaviour. These aesthetic values may count in favour of certain protections, and for that reason we will set these aesthetic values aside until section 3.3.

The three benefits addressed here—hunting assistance, transportation, and entertainment—share the structural feature that they require the animals to act in certain ways for the benefits to be derived: we need the animals to hunt for us, or to transport us, or to perform for us. Insofar as these benefits rely on animals *doing* certain things, they count in favour of certain basic rights: service animals have rights to the nutrition, housing, health-care, and security necessary for the animals to perform the relevant service. If these needs are not met, the animals cannot perform the services in question.

Like security, service can count for or against further rights depending on precisely what treatment the animals require to persuade them to perform the desired service. If punishment proves the most effective means to persuade an animal to hunt for you, to transport you safely where you please, or to entertain you or others in the desired way, the service in question counts against further rights for the animal. In that case, you can get what you want without needing to respect the interests of the animal in any given way. If, on the other hand, kindness is needed to persuade an animal to perform the desired service reliably and safely, then to that extent the benefits from

that service count in favour of rights.⁶³ Once more, a process of negotiation between the owners or handlers and the animals in question will be needed to determine precisely what further rights will be needed in order for the animals to perform the desired service.

Medical/Testing

The various ways that we use animals for roughly ‘medical’ ends can be divided into two groups: ‘medical service’ requires the animal to behave in certain ways to fulfill the function, while ‘medical testing’ does not require any specific behaviour on the part of the animal. To identify examples of medical service, consider our use of seeing-eye dogs to conduct the visually impaired safely about their environments, and of various companionship animals to help treat some mental conditions like depression or autism. To identify examples of medical testing, consider our use of a host of animals in the testing of drugs, treatments, and cosmetics.

The category of medical service is essentially a sub-category of service and thus shares the same governing principle: such a benefit counts in favour of basic rights to the conditions needed for the animals to perform any service at all. The benefit may count for or against further rights depending on the kind of treatment needed to get the animals to perform the desired service. If punishment and intimidation are the best ways to bring about the desired benefit, then medical service counts against further rights. On the other hand, if the animals will withhold their service unless further rights are recognized, then medical service counts in favour of whatever

63. Indeed, there is anecdotal evidence that the horses used for transportation purposes would go on strike in protest at their working conditions, showing that they require more than just punishment in exchange for the benefit of transportation. See Jason Hibril, “Animals, Agency, and Class: Writing the History of Animals from Below,” *Human Ecology Review* 14, no. 1, 101–112, discussed in Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 115.

further rights the animals and their owners or handlers agree upon.

The category of medical testing counts against rights: often no kindness is needed to get the benefit from the animals, and the benefit can often be derived only by confining the animals in controlled environments and causing them serious pain or death.⁶⁴

Nourishment

We use animals for nourishment in two different ways, either by eating the animals, or by eating something that the animals produce, like eggs or milk. For the animals we eat, the benefit of nourishment may count in favour of minimal rights: the benefit of nourishment will only be derived if the animals receive sufficient food, health care, and protection from the elements and predators. Thus, animals used for nourishment are entitled to the conditions necessary for them provide the benefit in question while they are alive. Their natures are such that, without sufficient food, protection, and health care, they will not grow to be optimal sources of nourishment, so although they cannot *deliberately* withhold benefits in retaliation for ill treatment, they are bound to do so by their natures. But the nourishment derived from animals we eat necessarily counts against a right to life: should we be forbidden from killing them, the benefit of nourishment would dissolve and it would no longer ground any rights for the animals in question.

Animals who produce something that we eat are in a similar situation: they have

64. Christine Sypnowich has noted, in private correspondence, that it is a striking weakness of the contractarian approach that the rights claims it generates are weakest in cases where animals are most vulnerable, for instance, in scientific experiments. Chapter 5 presents my full defence of the contractarian account, but it is worth acknowledging here that the weakness Sypnowich identifies is indeed a weakness of the contractarian account of animal rights. I will argue in chapter 5, however, that the contractarian account is still the best account from a security perspective.

rights to the conditions needed for them to provide the desired produce. Often, this will entail the same rights to adequate food, health care, and protection.⁶⁵

Clothing and Other Products

Many animals are used in the making of clothes and a variety of other products. In a small number of cases, these products do not require the killing of the animal in question, as in the case of sheep's wool. But in most cases, the animals must die for humans to get the products they want. In the former cases, the benefits derived can count in favour of rights for the animal in question to the extent that respecting some of the interests of the animal is instrumental to deriving the benefit in question: sheep require adequate protection and nourishment to yield wool. In the latter cases, the benefits count against rights for the animal in question: the animal must die for the benefits to be derived, and no amount of human kindness will improve or increase the benefits derived.

Hunting

Some humans hunt wild animals, whether for sport, nourishment, or both. The benefit humans reap from these activities come directly at the expense of animals, and no human restraint can enhance the benefit derived. Thus, these benefits count against rights for animals.

65. However, the right to "adequate" food is complicated when the interests of the animals and those of their human owners diverge. See footnote 68 on page 125.

3.2.2 Species

Section 3.2.1 has categorized animals by use and detailed the rights entailed by each use. This section will categorize animals by species, and detail the rights owed to various species. Of course, it is impossible to examine *each and every* species in this section; I will focus instead on dogs, cows, bears, and squirrels as species representative of the larger classes of, respectively, domesticated animals, farm animals, wild animals, and ‘liminal’ animals.⁶⁶ But once again generalizations, whether based on species or on use, will not be completely satisfactory as they are vulnerable to the many variations at the individual level; generalizing from these examples requires careful attention to the principles outlined in section 3.1.5.

Dogs

Dogs, in general, are capable of benefiting humans by providing companionship and security; by helping to hunt, transport, and entertain humans; by serving as subjects in medical experiments; and by providing medical services like guiding the blind. But to be certain, no one dog is capable of providing all of these benefits, and any given dog’s rights will depend on the benefits it provides. Most dogs are also capable of causing rather serious harms to humans. Certainly, well-armed humans need not fear dogs, but most of us are not well-armed. Depending on the dog, it could pose a lethal, moderate, or negligible threat to an adult human, and most could pose a moderate or lethal threat to a human child or infant. While human children and

66. While I am not following their categorizations precisely with these examples, the examples were chosen in light of Donaldson and Kymlicka’s categorization of animals into three groups: domesticated, wild, and liminal. My examination could not stay at such a general level, requiring particular examples and a further subdivision of the domestic category. Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*.

infants likely do not qualify for rights, for reasons discussed in section 3.1.5, those humans who do qualify for rights generally desire that their children should not come to harm. Thus, harm to human children or infants will cause harm to able adult humans. Moreover, most dogs are capable of withholding benefits in response to ill treatment, and of withholding harms in response to good treatment. They are also capable of communicating their interests to humans, and cooperating with humans.⁶⁷

Given this combination of abilities, dogs used for companionship, security, hunting assistance, transportation, entertainment, and medical service should qualify directly for some moral rights. Should one treat a dog without restraint, it would be capable of withholding significant benefits, and perhaps of imposing moderate harms. Thus, the benefits one would have to derive from disregarding the interests of the dog would have to be significant to make that course of action advantageous. Of course, not all of the dog's interests will call for respect: our interests in confining them, for instance, may be rather great, and the dog may not feel that this confinement is so offensive as to retaliate. So while a full list of dog rights will be quite controversial, and unlikely to hold for all dogs, certain basic rights should be rather uncontroversial. Dogs should have rights against physical and mental abuse, and rights to adequate nourishment, health care, and protection. To deprive a dog of nourishment, or to abuse it, is to turn that dog into a threat to all humans: dogs will not retaliate against their abuser alone, but will often be violent to all humans. Given that human adults care greatly about their children, and that dogs could pose lethal threats to their children, humans should find it advantageous to require that no dogs be abused or starved. We could eliminate these potential threats by simply killing all dogs, but that would deprive all humans of the chance at any of the benefits that dogs are capable of giving. And

67. See Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 104–105, 108–112.

without adequate health care, food, and protection, we will be unable to derive the benefits desired from dogs. Thus, it would be in everyone's interests to observe these restraints to ensure that dogs do not become threats.

One important thing to note about this approach to rights for dogs is that individual dogs have a role to play in determining their rights: dogs *take an interest* in things, and different dogs take an interest in different things. In determining what rights a dog has, it is not sufficient to consider what treatments the dog might have a reason to retaliate against; one must pay attention to what treatments the dog actually retaliates against. The "must" here is prudential: since one wants to derive benefits from a particular dog and avoid any costs that it may impose in retaliation, one ought to be concerned to avoid actual retaliation. It will still be useful to consider the interests that dogs generally tend to have insofar as this helps one to predict when one's particular dog will retaliate, but adjustments need to be made based on the particular interests that one's dog takes.

What about the use of dogs in medical experiments? Foreseeable benefits of experimentation may count against rights for dogs used for that purpose. In order for it to do so, the benefits to be derived from medical experimentation on dogs would have to outweigh all of the foregone benefits and the accompanying increased risk of harm caused by the dog. Perhaps the benefits of medical experiments can outweigh these foregone benefits, and perhaps the increased risk of harm can be neutralized by effectively confining the dog, and in that case there would be no way to justify a duty against testing on the dog. But the benefits of each particular experiment must outweigh the costs, so while dogs may not have a right against all medical testing, they should have a right against testing that did not have expected benefits sufficient

to outweigh the costs. And because some medical experiments may require abuse or starvation, dogs used for medical experiments lack rights against such treatment.

Thus, as a basic set of rights, dogs not used for medical testing would have the right not to be physically or mentally abused, and the rights to adequate nourishment, protection, and health care. All dogs would have the right not to be used in medical tests whose expected costs outweighed their expected benefits. But dogs used for appropriate medical tests would have no rights. Some dogs could have more rights on top of this, but those rights would vary more with the particular circumstances of the dog. For example, it would seem safe to say that a sled dog is entitled to protection against wild predators, insofar as its owner is capable of providing it. But these particular cases are too varied to be covered in this article.

Cows

Cows are capable of providing nourishment—both by being eaten and by providing their milk—and of providing clothing and other accessories. While cows are incapable of deliberately withholding benefits, their natures are such that, to provide some of the benefits we desire, they must be treated in certain ways. For instance, cows require adequate nutrition to yield milk or to be of use for eating. Moreover, they are highly dependent on human protection from predators and disease. Thus cows, and most farm animals, would only qualify for rights to those conditions needed for them to

provide any benefit to humans: adequate nourishment,⁶⁸ protection from predators, and health-care. However, these rights are conditional: should circumstances make it the case that the benefits of violating one of these rights outweigh the costs then the right in question would no longer hold. For instance, in a drought or famine, resources and energies might be better spent otherwise than by raising animals for food, in which case the rights of farm animals to adequate nutrition would no longer hold. Thus, farm animals would have rights to the conditions needed for them to provide the benefits we desire, yet those rights would be conditional on the benefits outweighing the costs of guaranteeing the rights in question.⁶⁹ Of course, this conditionality is not unique to animal rights, but a general feature of contractarian rights: should conditions change so significantly that the circumstances of justice no longer obtain, human rights would equally dissolve.

Bears

Humans derive benefits from bears by hunting them or by using them to make clothing or other accessories. Unlike animals who are farmed for clothing or other accessories, bears fend for themselves in the wild, and thus do not get rights to adequate food,

68. The right to adequate nourishment applies to the kind of nourishment that is both in the cow's interest and in the interest of the farmer. The primary case that I have in mind is undernourishment, where a farmer feeds a cow so little that the cow does not produce. The nourishment must be in the interest of the cow in order for the requirement to count as a moral right in accordance with the definition provided in section 1.1. The nourishment must be in the interest of the farmer in order for the farmer to have a duty—grounded in mutual advantage—to provide the nourishment in question. But as professor Kymlicka has noted in private correspondence, these interests can diverge: the sort of nourishment that maximizes the farmer's interest may not be the sort of nourishment that is in the cow's interest. In such a case, there is no room for mutual advantage, and the cow does not have a right against the nourishment that maximizes the farmer's interest. At the same time, by feeding the cow in this manner, the farmer cannot claim to be fulfilling a right of the cow.

69. One interesting corollary of this approach is that, should we all stop eating farm animals and their produce, the direct rights of cows and other farm animals would dissolve, as we would no longer reap a benefit from them that was conditional on certain treatment. Recognition of such rights could no longer be mutually advantageous.

protection, and health care from that benefit. Moreover, because bears are capable of imposing considerable harm on humans, and are incapable of refraining from harming humans in exchange for restrained treatment, they can be considered a threat and thus inadmissible for rights. So bears, and most wild animals, would not qualify for rights.

However, while one's risk of being attacked by a bear may never be brought to zero, no matter how much restraint one observes, there are certainly behaviours that will increase one's risk of bear-attack. Coming between a mother-bear and her cubs stands out as a particularly good example of behaviour that is bound to bring on attack. Given the great harms we are likely to suffer, and the minimal benefits one could derive from so positioning oneself, it would seem rational to avoid coming between a mother-bear and her cubs. Bears would have rights against other such behaviour.

Thus bears, and most wild animals, would not qualify for many rights, though they would have rights against any behaviour that would provoke them to attack a human.

Squirrels

Squirrels belong to a class of animals that Donaldson and Kymlicka call "liminal animals."⁷⁰ They are not domesticated animals, but they seek out human societies in which to live, in contrast to wild animals. Squirrels, and liminal animals more generally, provide no benefits to humans, and what minor harms they cause humans they cannot avoid causing. Thus, such animals would not qualify for rights directly.

70. Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 210–251.

3.2.3 Summary

This section has applied the criteria from section 3.1.5 to animals categorized by use and by species. Based on the first set of categories, animals used for companionship, security, hunting assistance, transportation, entertainment, medical service, nourishment, and clothing or other accessories are owed basic rights to adequate food, health care, and protection from predators and the elements. Animals used for companionship are owed further rights, which must be negotiated on a case by case basis between the animals and their owners. Animals used for security, hunting assistance, transportation, entertainment, and medical service may be owed further rights if they withhold the benefit in question unless certain additional conditions are met, and if those conditions are not more costly than the benefit derived. Animals used for medical testing or hunting have no rights: the benefit derived requires that the animals' interests be disregarded. All of these rights are conditional on the animal continuing to provide the benefit in question: should the animal be transferred from one category to another, or cease to provide benefit altogether, the rights of that animal may change or cease to exist entirely.

On the basis of the categorization according to species, dogs should have rights against mental and physical abuse and against use in 'frivolous' medical experiments, and rights to adequate nourishment, health care, and protection from predators and the elements, unless they are serving in medical experiments. Dogs may also receive further rights, as negotiated between dogs and their owners. Cows should have rights to adequate food, health care, and protection, but bears and squirrels do not receive any rights. We may, with caution, generalize from dogs to domesticated animals, from cows to farm animals, from bears to wild animals, and from squirrels to liminal

animals.

These two sets of categorizations help to offer guidance on the rights owed to general categories of animals, but because there is great variation in the abilities of individual animals within each category, these guidelines must be adjusted in light of the principles detailed in section 3.1.5 and further information about the abilities and interests of individual animals.

3.3 Protectorate Status

For an individual to qualify indirectly for protectorate status, either some set of people who qualify directly for rights must feel strongly enough that the individual in question ought not to be treated in certain ways that they would be willing to carry out threats against anyone who treats the individual in that way, or else all those who qualify for rights must decide that it is in their interest to obey certain restraints that protect the individual in question. In the former case, those forming the coalition find any agreement with the excluded group to be less advantageous than that with the current coalition, and thus coerce the excluded group to comply with the protections in question. The latter case amounts to an insurance-based protection, and need not rely on coercion. As explained in section 3.1.2, threat-based protections will be rare because they are costly, and will only emerge when the excluded group is very small, weak and disorganized; when the proscribed action is easy to detect; when the costs of detection and punishment are low; when the precedent effects can be minimized; and when the foregone benefits of a broader coalition are least. Thus, we will only tend to enact protections that protect our most important interests, like our interests in the wellbeing of our children.

The first protection that is likely to be established is a threat-based protection for pets against non-owners. Many people enjoy keeping animals as pets, and find themselves in a similar situation toward their pets as they would be toward their infant children: they take a deep interest in the well-being of their pets, and any harm to their pets harms them.⁷¹ And the vast majority of people take little interest in harming others' pets. Insofar as pet owners would likely carry out some retaliation against anyone who harmed their pets, and no one really has a strong reason to abuse another's pet anyway, a convention recognizing that pets should not be harmed would likely emerge. But it is important to qualify this convention: the purported duties not to harm would apply to all but the pet's owner. It says nothing about a pet's owner abusing it. This protection is akin to the protections extended to a person's property: no one is permitted to destroy my car, but I am free to do so.

Would any protectorate status be recognized that included protections for animals against their owners or, should the animals be un-owned, against all humans? Here we would have to make reference to the preferences that people have over wild animals, liminal animals, and domesticated animals that are not owned by the people in question. A considerable majority of people feel rather disgusted when they are made aware of any cruelty inflicted on animals, where cruelty is understood as violence that is disproportional to, or unnecessary for, whatever aim is thereby achieved. And there is a growing trend toward vegetarianism and veganism, involving disgust at any harm to animals, or even use of animals as mere means toward human ends. Nevertheless, it seems that the vast majority of people are still fine with the killing and harming of farm animals, wild animals, and liminal animals, so long as some benefit is thereby obtained and the harm is necessary to obtain the benefit. Thus all animals could

71. Narveson, "On a Case for Animal Rights," 44.

have protections against cruelty, or harm that is disproportional or unnecessary to any benefit thereby obtained, but they would lack stronger protections against any harm or against being used as mere means.

Aside from our love of pets and our discomfort with animal cruelty, many also have aesthetic preferences regarding nature: we want natural environments to be preserved, and we want the natural diversity of species preserved for our appreciation and for the appreciation of future generations. The strength and prevalence of this preference will determine the extent to which some wild animals might receive protections from any harmful or disruptive human interaction in their habitats. These preferences may not protect all or even many wild animals, however: the preference might be satisfied by maintaining a certain number of natural habitats and animals, while allowing others to be destroyed. It will depend upon the specifics of the preference, whether people prefer that all wild animals be allowed to thrive in their current environments, perhaps with positive human interventions to correct for the detriments already imposed by human development, or whether they simply want a sufficient diversity of wild habitats to be preserved for human enjoyment. Either way, some limited protections could be established based on our aesthetic preferences for the preservation of nature.

Finally, nearly all humans strongly prefer to take actions today to secure their future well-being, and to avoid actions that would imperil their futures. Moreover, most humans care strongly about the future well-being of their offspring. As such, climate change should be of great concern to all humans, and more generally, all humans should strongly desire to preserve the ecosystems necessary to support and promote human life.⁷² Insofar as the disruption of certain ecosystems threatens the ability of the earth to support human life, strong protections for the animals requisite

72. Narveson, "On a Case for Animal Rights," 45–46.

for the health of those ecosystems will be generated, and these protections should hold based on insurance reasoning. Again, not all wild animals will be protected equally, but at the very least some wild animals will receive rather strong protections from negative human interference in their lives and habitats, and will perhaps even be owed positive assistance from humans, such as inoculation against disease.

To sum up, rather strong and universal protections can be generated to protect those wild animals required for the continuation of human life, including those animals and ecosystems whose destruction would contribute to climate change. These protections would include protections against negative human interference, and likely an entitlement to beneficial human intervention. Perhaps only a bit less universally, protections for all animals against cruelty, or harm disproportional to whatever benefit is to be gained thereby, are likely to arise from human discomfort at the thought of cruelty to animals. Less universally still, protections for pets against harm at the hands of anyone other than their owner are generated by humans' love of their pets. And likely the least universal and the least strong of all, aesthetic preferences for the preservation of certain wild spaces and species would lead to protections against negative human intervention and perhaps even entitlements to beneficial human intervention for some wild animals. I say least 'strong' because our aesthetic preferences are generally rather weak, and we can be compelled to ignore our aesthetic preferences for the sake of some stronger preferences. Thus, poorer societies, while they may have the same aesthetic preferences for natural preservation, are likely to be driven to exploit their natural resources in whatever way they can to fulfill their stronger preferences for adequate food, shelter, and medical resources. Aesthetic preferences may generate protections in wealthy societies where natural preservation can be achieved

without sacrificing more urgent needs, but will fail to do so elsewhere. The same is true of all of protections: should the human desires that these protections satisfy be overwhelmed by more urgent desires, the fulfillment of which required the violation of those purported protections, then those protections would cease to hold.

3.4 Objections

While the revised contractarian theory developed here is better able to ground rights for animals and marginal cases, many may still find it inadequate because it violates their moral intuitions concerning animals and marginal cases. Chapter 4 will examine the most prominent alternative accounts of rights which yield more generous rights for animals and marginal cases. And while chapter 5 will present my full argument that social contract theory offers a better theory of rights than the alternatives, there are two standard objections to the contractarian theory that are prominent in the animal rights literature that I should address here: the argument from marginal cases, and the “conflation” objection. I hope to show that the contract theory developed here is immune to these objections, and thus even if the theory is deeply counter-intuitive, the theory is consistent and deserves to be taken seriously.

3.4.1 Argument from Marginal Cases

A prominent argument that finds articulation both as a negative and a positive argument is known as the “argument from marginal cases.”⁷³ In its simplest negative form, the argument from marginal cases claims that any theory that guarantees rights for all and only humans is either inconsistent or arbitrary. To ground such rights, a

⁷³ Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 26–27; Singer, *Animal Liberation*, 16, 18–20.

theory will cite some natural property like language or reason and claim that possession of this property is a necessary and sufficient condition for being a right-holder. The argument from marginal cases contends any such natural property will either include all humans and some animals or else exclude all animals and some humans—the “marginal cases,” typically comprising humans who are disabled, young, or elderly. Thus, such theories must appeal to the natural property inconsistently in order to get rights for all and only humans, or else they must simply take “humanity” as the natural property, making such theories arbitrary and “speciesist.”⁷⁴

The argument from marginal cases leads us to a dilemma: either we must extend moral protections to animals, or else we must accept that some humans will be denied moral protection.⁷⁵ Of course, the philosophers who offer the argument from marginal cases are not indifferent about which option we choose, but in order to eliminate one horn of the dilemma one must appeal to positive arguments that animals or disabled humans ought to receive moral protection. In chapter 4 we will examine the various positive arguments offered by different philosophers.⁷⁶ But it is worth noting that the argument from marginal cases ties together the ‘moral fates’ of animals, human infants, and humans with certain severe disabilities, motivating this thesis to consider these groups together.

In its negative form as an argument centering around inconsistency, the argument

74. Singer, *Animal Liberation*, 18–20.

75. Singer explicitly acknowledges that the argument from marginal cases could lead us to either of these two conclusions. *ibid.*, 19–20.

76. Some philosophers try to eschew the need for a positive argument, simply claiming that excluding marginal cases from moral protection is obviously absurd. By claiming that one horn leads to absurdity, such philosophers force us to extend moral consideration to animals as well as marginal cases. But such claims of obvious absurdity are merely attempts at intellectual bullying, and fail to persuade anyone who does not already agree. See, for instance, David DeGrazia, *Animal Rights: A Very Short Introduction* (Oxford: Oxford University Press, 2002), 17, 25–26; Clare Palmer, *Animal Ethics in Context* (New York: Columbia University Press, 2010), 21–22.

from marginal cases does not apply to the kind of mutual advantage social contract theory that I have developed in chapters 2 and 3.⁷⁷ As we saw, to the extent that animals and marginal cases are excluded from moral protection, it was because able humans would not find it to be in their interest to restrain their behaviour toward such individuals. Excluded animals and marginal cases, it was claimed, are too weak to withhold any benefit or to pose any threat in retaliation for unrestrained treatment, or are incapable of responding properly to human behaviour. Insofar as such a mutual advantage theory applies the criteria for moral rights and protections consistently, treating animals and marginal cases in the same manner as able adult humans, it does not violate the argument from marginal cases in its negative form.⁷⁸

The arbitrariness of the mutual advantage criterion is still an open question. The theory is not “speciesist,” but one may still object that the natural properties that ground the possession of rights according to mutual advantage theories are morally irrelevant or arbitrary. Chapter 5 will argue that mutual advantage social contract theory offers the strongest account of rights, and in so doing aims to show that the criterion is not arbitrary. Thus the argument from marginal cases does not provide an objection to such contract theories. This shows that the argument from marginal cases does not pose a theory-independent objection to mutual advantage contractarianism, and thus its force, when applied to contractarianism, is entirely dependent on some

77. This is not meant as a criticism of those who put forward the argument from marginal cases. The argument is typically brought against those who endorse “utilitarianism for animals, Kantianism for people,” where the argument is quite pressing (Nozick, *Anarchy, State, and Utopia*, 39). Because many people find the argument from marginal cases to be particularly powerful, it is important to show that the argument does not count against the kind of social contract theory defended in this thesis.

78. Even the unrevised contract theory discussed in chapter 2, which applied the wrong criteria for moral rights, nevertheless did so consistently. The mistake was not one of inconsistent application of a criterion, but rather of an incomplete grasp of the natural properties grounding rights for contract theory.

ethical theory's claim to show that mutual advantage is an inappropriate basis for moral rights. And it is the latter claim that chapter 5 will attempt to refute.

3.4.2 Conflation

Martha Nussbaum has argued that social contract theorists, including Hobbes, have conflated two questions that could and should have been kept distinct: "Who frames the principles of justice?" and "For whom are these principles framed?"⁷⁹ There are two ways that one could interpret "conflate" in this objection: either it means that the social contract theorists mistakenly collapsed the two questions, for instance by failing to appreciate that the two questions were distinct, or it means that the social contract theorists argued or believed that the two questions were connected for some reason. On neither interpretation is Nussbaum's objection compelling.

The first interpretation, involving a failure to appreciate the difference between the two questions, fits best with what Nussbaum actually says. For instance, Nussbaum claims that "the classical theorists all *assumed* that their contracting agents were men who were roughly equal in capacity, and capable of productive economic activity," that "the classical social contract thinkers *failed to imagine* them [the physically and mentally impaired] as participants in the choice of political principles," and that "the contracting parties *are imagined* to be one and the same as the citizens who will live together and whose lives will be regulated by the principles that are chosen."⁸⁰ But on this interpretation, Nussbaum's claim is descriptively false. Because Hobbes and Hume (in his mutual advantage writings) contended that the only way to ground moral obligations was to show that they were in the interest of the obligated parties,

79. Nussbaum, *Frontiers of justice*, 21.

80. My emphasis. *ibid.*, 14–16.

this means that by definition any moral obligations will be in the interest of the obligated parties—the parties ‘framing’ the principles of justice. This is not a failure to distinguish two questions; it is an argument that the questions are connected. Moreover, it is not true that Hobbes or Hume failed to imagine animals or marginal cases as right-holders. In chapter 2 we saw that both Hobbes and Hume *explicitly* considered and refuted the possibility of animals and marginal cases as right-holders based on their theoretical commitment to mutual advantage as the purpose of justice. Animals and marginal cases, it was claimed, are too weak to withhold benefits or to mete out punishments that would be sufficiently impressive to make it in the interest of able humans to restrain their behaviour toward them. If we take Nussbaum’s objection to be that contract theorists failed to consider the possibility that animals or marginal cases could be right-holders, or that non-parties to the contract could be right-holders, it is glaringly false. Moreover, the contract argument developed in this chapter has not conflated these questions, so even if Nussbaum’s objection applied to Hobbes and Hume, it should not apply to the contract theory developed in this chapter.

The second interpretation, involving a reasoned connection between the two questions, fits less well with Nussbaum’s claims. Indeed, on this interpretation, Nussbaum’s claim that contract theorists conflated the two questions would no longer amount to an objection but instead to a mere observation. This interpretation takes into account that contract theorists offered arguments connecting the two questions: if we are to ground rights in mutual advantage, then only the interests of duty-holders will guide the articulation of moral principles. But to make this an objection rather

than a mere observation requires appeal to a positive argument that mutual advantage is an inappropriate basis for moral concern, or that another basis offers more desirable principles. One could take Nussbaum's proposal for grounding rights in capability theory as a refutation of the claim that only mutual advantage will do to ground rights, but then this objection rests on the success of Nussbaum's capability theory. Section 4.3 will examine the details of Nussbaum's capability theory, and chapter 5 will argue that the capability theory offers a weaker account of rights than contractarianism.⁸¹ For now, it is enough to show that conflation is not an objection on its own, but rests entirely on the the success of Nussbaum's capability theory, to show that this objection does not on its own pose a challenge to the contract theory developed here.

3.5 Conclusion

I hope to have shown that social contract theories rooted in mutual advantage do indeed have the resources to generate moral rights and protections for animals, contrary to the views of both critics and advocates of social contract theory. According to mutual advantage contractarianism, animals used for companionship, security, hunting assistance, transportation, entertainment, medical service, nourishment, or clothing tend to qualify for rights against starvation, predation, and disease. Dogs tend to qualify for rights against abuse and frivolous experimentation. Cows tend to qualify for rights against starvation, predation, and disease, but squirrels and bears tend not to qualify for rights. These tendencies offer us guidance with respect to our treatment of animals, but this guidance often needs to be supplemented with information about

81. Section 4.3 is on page 149, and chapter 5 is on page 161.

the individual animal in question. Mutual advantage contractarianism also recognizes potential protectorate status for various animals. All animals have certain protections against cruelty stemming from protectorate status. Moreover pets and those animals needed for the environment to support human life or to fulfill our aesthetic preferences would also receive such moral protections.

These rights and protections are certainly not as comprehensive as most animal rights advocates would like, but they are moral protections nonetheless. The standard quick dismissal of such social contract theories as inadequate to the task of animal rights is thus mistaken; social contract theories deserve a fair hearing. Moreover, by arguing that mutual advantage contract theory is not open to the argument from marginal cases or the conflation objections, I hope to have shown that contract theory is consistent and coherent. We must compare the strengths and weaknesses of social contract theory with those of the alternative theories preferred by animal rights theorists, who ground their claims in capabilities theory, deontology, and utilitarianism. To this end, chapter 4 will examine the details of the dominant theories of animal rights, and chapter 5 will compare these theories to the contract theory, and argue that the latter offers us the strongest account of rights from a security perspective.

Chapter 4

Dominant Theories of Animal Rights

In this section I will examine the leading alternative moral theories which claim to do a better job of protecting the rights of animals and marginal cases.¹ The theories that I will examine are Singer’s utilitarian account of animal justice, Regan’s Kantian account of animal rights, and Nussbaum’s capability theory of rights for animals and marginal cases. In examining these theories, I will pay particular attention to the foundations that they offer for their claims—their answers to the questions “Why is this the correct moral theory?” and “Why should I behave as this theory says I ought to?” The primary purpose of this section is to explain these alternative theories and their foundations; I will save my main critique of these arguments for chapter 5.²

4.1 Utilitarian Theory of Rights

In *Animal Liberation*, Peter Singer offers an argument that animals, and marginal cases in general, are owed equal moral consideration on utilitarian grounds.³ Based

1. Though it is slightly funny for a mutual advantage contract theorist to call these theories “alternatives” as they are currently much more popular than mutual advantage theories.

2. Page 161.

3. Singer, *Animal Liberation*.

on the way that this thesis is using the term “moral right,” Singer’s claim is equivalent to the claim that animals and marginal cases have a basic moral right to have their utility weighed equally with that of others in utilitarian calculations.⁴ His argument rests on three key propositions. The first is “the basic principle of equality,” which holds that the interests of all beings ought to receive equal consideration⁵. The second I will call “the sentience boundary,” which claims that sentience—the ability to feel pain and pleasure—is both a necessary and a sufficient condition for a being to have any interests at all, and thus to have a right to equal moral consideration.⁶

If a being suffers there can be no moral justification for refusing to take that suffering into consideration. No matter what the nature of the being, the principle of equality requires that its suffering be counted equally with the like suffering—insofar as rough comparisons can be made—of any other being. If a being is not capable of suffering, or of experiencing enjoyment or happiness, there is nothing to be taken into account. So the limit of sentience (using the term as a convenient if not strictly accurate shorthand for the capacity to suffer and/or experience enjoyment) is the only defensible boundary of concern for the interests of others.⁷

To mark the boundary of moral consideration any other way, Singer claims, would be arbitrary.⁸

4. See section 1.1 for an explanation of the sense of basic moral rights.

5. Singer, *Animal Liberation*, 2-5.

6. *ibid.*, 7-8.

7. *ibid.*, 8-9.

8. *ibid.*, 8-9.

To the basic principle of equality and the sentience boundary, Singer adds the duty to minimize the suffering of all sentient beings.⁹ Thus, we all have a duty to act so as to minimize the expected suffering of sentient beings whose suffering is weighed equally.

Because animals and marginal cases are sentient, it follows from the first three premises that their interests ought to receive moral consideration equal to the interests of every other sentient being. So, for example, if a father is trying to decide whether it is acceptable to make his daughter her favourite meal—duck stew—on her birthday, he ought to consider the happiness that his daughter gets from eating duck stew and the suffering that the duck must go through, and compare the net balance of pleasure and pain associated with the duck stew to the net balance of pleasure and pain associated with all of the other possible meals he could serve. For the sake of simplicity, let's suppose that the only other alternative is to serve a lovely vegan meal, from which his daughter may not derive much pleasure, but which will not cause any suffering to any animals (so long as we bracket all sorts of considerations about how the ingredients were transported from their place of origin to the dining table, etc.). Bracketing complications associated with inter-species comparisons of utility, it seems reasonable enough to say that serving the vegan meal will produce the highest net utility, and thus the father ought not to serve duck stew. For the father to serve duck stew would be for him to discount or ignore the interests of the duck, in violation of the basic principle of equality.

Singer examines some of the implications this view has on the ways we currently use animals. Firstly, Singer examines animal medical testing. Utilitarianism does not

9. While Singer uses this negative formulation, utilitarianism generally appeals to the prescription to maximize expected utility, and it seems likely that the two formulations should be interpreted as equivalent. Singer, *Animal Liberation*, 17.

rule out all animal testing: Singer believes that there may be circumstances where a test could be performed that caused suffering to one individual (or a few individuals) but thereby produced a treatment that alleviated the suffering of many individuals. Such a test would minimize aggregate suffering.¹⁰ However, our current practice of preferring to experiment on animals rather than humans strictly based on species grounds is indefensible: if we are willing to experiment on animals, we must be willing to experiment on marginal cases and other humans, choosing those tests and those test subjects based on considerations of expected utility, not of species.¹¹ In effect, while utilitarianism does not proscribe animal experimentation, it would nonetheless determine that many of the experiments we currently perform cause greater aggregate suffering than would different tests (or no tests), and are thus impermissible.¹²

Secondly, Singer examines our factory farming practices. Such farming practices fail to minimize expected aggregate suffering, and are thus proscribed. Moreover, engaging in factory farming desensitizes humans and has the effect of leading humans to disregard animal suffering. Thus, factory farming leads us to reason immorally, disregarding the suffering of animals, and is thus proscribed also on those grounds.¹³

Thirdly, Singer examines our practice of eating animals. This practice causes immense suffering to animals, destroys the environment, desensitizes humans, and amounts to an inefficient use of resources for feeding humans, thereby causing human suffering. Moreover, for the most part, humans do not need to eat meat. Thus, in cases where eating animals is not needed to survive—or generally, where eating animals will not minimize aggregate expected suffering—it is proscribed.¹⁴ It should

10. Singer, *Animal Liberation*, 81–85.

11. *ibid.*

12. *ibid.*, 25–94.

13. *ibid.*, 95–157, 229.

14. *ibid.*, 159–183.

be noted, however, that Singer thinks utilitarianism proscribes the imposition of suffering, but does not proscribe painless killing.¹⁵

But what grounds does Singer offer for the three fundamental principles? Singer does not take special pains to defend the principle of equality in *Animal Liberation*. He notes that there is broad consensus that some formulation of this principle underlies all moral theories, but offers little defence aside from this appeal to popularity.¹⁶ So his arguments appear to be targeting an audience that already accepts the basic principle of equality as self-evident. In *Practical Ethics*, Singer offers the following argument for the basic principle of equality and for the duty to minimize suffering. The argument starts with the claim that the most basic component of ethics and of living ethically is to attempt to justify one's actions to others.¹⁷ Such justification requires a certain level of universalism: I cannot justify my behaviour by reference to my own preferences alone. To justify my behaviour to others, I must adopt a relevantly universal perspective that they can share.¹⁸ Singer proposes that, to universalize in the relevant way, I should take up, in turn, the perspectives of those to whom I am justifying myself, and evaluate my actions based on their preferences. Once I attempt this evaluation, I see that to justify universally, I cannot prefer my interests over the interests of others, and thus the basic principle of equality is derived. It takes little further work to add the duty to minimize suffering: when I am taking up the various perspectives and considering their preferences, it seems straightforward that (a) that each individual wants to minimize his or her own suffering, and (b) that each individual wants to satisfy as many of his or her preferences as possible. Thus,

15. Singer, *Animal Liberation*, 20–21.

16. *ibid.*, 5.

17. Peter Singer, *Practical Ethics*, Third (Cambridge: Cambri, 2011), 9.

18. *ibid.*, 9–10.

to justify my actions, I must show that they minimize aggregate suffering, or (what may be equivalent), that they maximize expected aggregate preference satisfaction.¹⁹ Thus, Singer claims that we derive utilitarianism by universalizing non-ethical egoistic expected utility maximization.²⁰

Singer's argument can be found throughout the utilitarian tradition in various forms. For instance, in John Stuart Mill's *Utilitarianism*, Mill claims that each human values his or her happiness, and that any other interest he or she has is an interest only insofar as it is instrumental to his or her happiness.²¹ The argument then infers that since "each person's happiness is a good to that person, . . . the general happiness, therefore, [is] a good to the aggregate of all persons."²² While this "aggregate of all persons" may at first appear unrelated to any individual human, they are connected by the emphasis on universalism, and the aggregate of all person can be equated to Sidgwick's "point of view of . . . the universe,"²³ or Hume's "judicious spectator,"²⁴ all of which designate a universal impartial perspective from which actions are to be justified. We need to take up these perspectives because, were we each to evaluate things from our own perspectives, our judgments would disagree. My own interests may lead me to judge differently than your interests lead you to judge. Insofar as we seek agreement in our moral judgments, we need to take up a shared perspective, and the only shared perspective that each of us could agree to is one wherein each of our interests are taken into equal account.²⁵

19. Singer, *Practical Ethics*, 11–12.

20. *ibid.*, 14.

21. John Stuart Mill, "Utilitarianism," in *Utilitarianism and On Liberty*, Second, ed. Mary Warnock (Malden, MA: Blackwell Publishing, 2003), 210.

22. *ibid.*

23. See book III, chapter xiii, section 3. Henry Sidgwick, *The Methods of Ethics*, Seventh (Indianapolis: Hackett Publishing Company, 1981), 382.

24. See III.iii.1. Hume, *A Treatise of Human Nature*, 632.

25. *ibid.*, 625–652 (III.iii.1–2).

Thus, according to Singer, all sentient beings have the right to equal moral consideration in the utilitarian calculus. Utilitarianism requires that we maximize expected aggregate utility and minimize expected aggregate suffering. We are thus duty-bound to promote animal utility, and to promote it equally with the promotion of human utility. We are not, however, proscribed from killing animals painlessly, or any other behaviour that does not impact a being's pain or pleasure.

4.2 Kantian Theory of Rights

In *The Case for Animal Rights*, Tom Regan offers a Kantian argument for animal rights. The argument is targeting individuals who endorse Kantianism-for-humans and Utilitarianism-for-animals, so for the most part my perspective is not the target of Regan's main arguments.²⁶ Regan begins with the premise that some beings have an inherent value that commands the respect of all other beings, or at least those rational beings capable of recognizing it.²⁷ In contrast to utility, this inherent value is distinct from the value of those individuals' experiences: an individual's inherent value does not vary with the quantity or quality of her experiences, nor with the quantity or quality of the experiences she causes for others. Moreover, inherent value is incommensurate with any other values.²⁸

Regan then argues that any beings who satisfy the "subject-of-a-life criterion" have this inherent moral value. Creatures satisfy the subject-of-a-life criterion if and

26. Regan does specifically critique the views of Jan Narveson, who endorses a Hobbesian view of morality where rights and duties are grounded in mutual advantage. So it is fair to say that Regan wants to address those of us who endorse justice as mutual advantage. Nevertheless, his arguments are targeted more toward the Kantianism-for-humans, Utilitarianism-for-animals view.

27. Tom Regan, *The Case for Animal Rights* (Berkeley: University of California Press, 1983), 233–235.

28. *ibid.*, 235–236.

only if they have:

beliefs and desires; perception, memory, and a sense of the future, including their own future; an emotional life together with feelings of pleasure and pain; preference- and welfare-interests; the ability to initiate action in pursuit of their desires and goals; a psychophysical identity over time; and an individual welfare in the sense that their experiential life fares well or ill for them, logically independently of their utility for others and logically independently of their being the object of anyone else's interests.²⁹

Thus, because many if not most animals and marginal cases satisfy the subject-of-a-life criterion, those beings have inherent value.

The argument then attempts to establish a "harm principle." Regan appeals to our considered judgments and claims that we have the considered judgments that:

It is wrong, other things being equal, to kill a moral agent, to make moral agents suffer, to deny moral agents access to opportunities for the satisfaction of their desires when it is in their interest to satisfy them. . . . Now, these beliefs, and others besides, share an important feature. They all involve prohibitions against harming moral agents. For harms. . . diminish an individual's welfare. . . . Because these beliefs share this common feature, it is possible to articulate a general principle (*the harm principle*) that unifies them. This principle states that *we have a direct prima facie duty not to harm individuals*.³⁰

This harm principle is only intended to protect individuals with inherent value. He

29. Regan, *The Case for Animal Rights*, 243.

30. *ibid.*, 187.

claims that it is fallacious to exclude from this principle moral patients—those agents who have inherent moral value, are incapable of controlling their actions, yet are capable of being harmed and wronged³¹—by arguing that moral patients cannot be harmed in the same way (or as badly) as moral agents can. To make such an argument is to treat the inherent value as reducible to, or commensurate with, the value of experiences.³² Thus, to exclude moral patients from moral consideration is arbitrary; such exclusion cannot be supported once one acknowledges that moral patients have equal inherent value, and Regan claims that we must deny the possibility that moral patients have varying degrees of inherent value if we want to avoid endorsing sexism, racism, or slavery.³³

Regan also articulates a further respect principle: principles of justice cannot discriminate between individuals who have equal inherent value, and must respect their inherent value.³⁴ Regan thinks that we fail to respect the equal inherent value of a being whenever we use it as a mere means, and thus such treatment is forbidden by the respect principle.³⁵

It may look like “inherent value” is playing a large role in Regan’s thought, but I think this is misleading. Regan’s arguments do not rely on such an appeal to objective values. When Regan talks about inherent value, this seems to be a short-hand for a group of prescriptive statements, so that the claim that “X has inherent value” is reducible to the claim that “X should not be treated in manners A, B, or C.” Regan grounds these prescriptive statements in our “considered judgments” and the

31. Regan, *The Case for Animal Rights*, 152.

32. *ibid.*, 187-189, 239-240.

33. *ibid.*, 236-237.

34. *ibid.*, 248.

35. *ibid.*, 249.

principles that best “explain” them.³⁶ Thus, to explain why anyone has “inherent value”—that is, why someone cannot be treated in certain ways—Regan appeals to our considered judgments that it is wrong to harm moral agents, and the harm principle that explains it. To show that inherent value is incommensurate with experience value, Regan argues that we intuitively judge the lives of two human agents with disparate experiences as equally deserving of certain treatments. And to show that all beings that have inherent value have it equally—that is, that it is equally impermissible to treat anyone in the proscribed manner—Regan appeals to our considered judgments that racism, sexism, and slavery—forms of arbitrary discrimination—are wrong.³⁷ Thus it appears unnecessary to interpret Regan as making claims about objective moral values, and given Regan’s endorsement of “Occam’s razor,” it seems most charitable to interpret his arguments in the least metaphysically demanding manner.³⁸

Thus, according to Regan’s view, many but not all animals and marginal cases are subjects of a life. All subjects of a life hold moral rights not to be harmed or to be disrespected. Harm is not restricted to causing pain: painlessly killing, causing suffering, or blocking the fulfillment of desires count as harm. The right not to be disrespected entails that we cannot arbitrarily discriminate between subjects of a life—for instance, by eating or testing on some but not others based on an arbitrary distinction like species—nor can we use subjects of a life as mere means. Thus, eating animals would be forbidden—even in cases where eating them was essential to our survival. While we are not bound to promote animal interests, we have stringent duties not to harm, use as mere means, or arbitrarily discriminate against animals

36. Regan, *The Case for Animal Rights*, 127–135, 146, 189–192.

37. *ibid.*, 235–240, 259–260, 267–268.

38. *ibid.*, 7, 146.

and marginal cases meeting the subject-of-a-life criterion.

4.3 Capabilities Theory of Rights

In *Frontiers of Justice: Disability, Nationality, Species Membership*, Martha Nussbaum articulates and defends a capability theory of rights for animals and humans with disabilities. In general, the capabilities approach specifies a list of capabilities that are deemed essential to a dignified life for a certain kind of being. Here, “capability” means something similar to substantive freedom: one has the capability to X if and only if it is feasible for one to do X given one’s current abilities and resources and the circumstances in which one finds oneself. For humans, the capability to do a specific set of things will be essential to a flourishing life, while for dogs, the ability to do a different set of things will be essential to a flourishing life, and so on for other species. In this way, capabilities are defined around a species-norm: to determine what capacities are needed for a dignified life for a certain individual, one must consider the species of that individual and the kinds of abilities that a member of that species needs to live well.³⁹ Unlike utilitarianism, the capability approach is non-aggregative: each capability on the list for a species is important, and losses in one capability cannot be compensated for by gains in another.⁴⁰ Nussbaum intends this as a partial political theory, intended to supplement Rawlsian political liberalism with a more complete set of fundamental rights defined by the list of capabilities for each species.⁴¹ Societies are then responsible for ensuring that individuals have the minimum set of capabilities required for a member of their species to live a dignified

39. Nussbaum, *Frontiers of justice*, 180–182, 362–369.

40. *ibid.*, 166–167.

41. *ibid.*, 23–25, 75, 155, 274, 284, 290.

life.⁴²

This brief sketch of the capability theory invites two important questions: first, how are the capability lists for each species to be defined, and second, why should we be concerned to guarantee the core capabilities for each individual? The two questions are closely connected, for any answer to the first question will rely heavily on the motive to which capability theory is supposed to appeal. If, for example, we were concerned for capabilities because of our concern for pleasure and pain, we would assume that the capabilities lists would be defined based on the pleasure that certain capabilities cause for individuals, and the pain individuals experience when those capabilities are lacking. In this way, we should expect the motive for caring about capabilities in the first place to play a fundamental role in deciding which capabilities are selected as being essential for members of a specific species. I stress the connection between the motive for recognizing the capability theory of justice and the values guiding the selection of capabilities because, as we will see, I think Nussbaum misrepresents the motive for endorsing the capability theory of justice. And in general, my attempts to identify Nussbaum's answers to these two questions in what follows are messy: Nussbaum draws alternately on Kant, Aristotle, Marx, Singer and Rawls to answer these two questions, yet it is not always clear that these ideas fit together well, or which ideas are meant to be grounding which arguments.

To determine what capabilities belong on a list of *core capabilities* for a species—on a list of capabilities that ought to be guaranteed as a matter of basic minimum

42. As discussed in section 1.1, while Nussbaum's theory is explicitly political, the sense in which she uses "political" overlaps with this thesis's focus on basic moral rights. I contend that we are discussing the same thing, even though she chooses to couch her claims in explicitly political terms.

justice—Nussbaum claims, quite rightly, that we cannot simply consult nature.⁴³ Creating a list of core capabilities for a species is an evaluative process, not a descriptive process. To test whether or not a given capability ought to be on the list of core capabilities for a given species, we are to imagine a life without that capability, and to ask ourselves whether or not such a life would be a dignified life for a member of that species.⁴⁴

Nussbaum distinguishes a Kantian notion of dignity from an Aristotelian notion of dignity, and claims that she is working with the latter. For Kant, “dignity” was a kind of non-fungible worth, contrasted with “value,” which was fungible: something with a certain value could be replaced by something else of equal value, whereas something with dignity has no equivalent and is thus irreplaceable.⁴⁵ And for Kant, a being had dignity only to the extent that it was capable of enacting his ultra-rationalistic morality, and thus only rational humans had dignity.⁴⁶ Nussbaum rejects this ultra-rationalistic conception of dignity in favour of an Aristotelian conception, which:

sees the rational as simply one aspect of the animal, and, at that, not the only one that is pertinent to a notion of truly human functioning. More generally, the capabilities approach sees the world as containing many different types of animal dignity, all of which deserve respect and even awe.⁴⁷

Here, Nussbaum seems to equate the term “dignity” with something we would more

43. Nussbaum, *Frontiers of justice*, 181–182, 366.

44. *ibid.*, 78, 352–356.

45. Kant, *Fundamental Principles of the Metaphysics of Morals*, 63–64.

46. *ibid.*, 64.

47. Nussbaum, *Frontiers of justice*, 159.

frequently call “flourishing,” or the idea of a good life for a particular kind of being.⁴⁸ This makes sense, given that it is not really clear what we are supposed to be imagining when we are to consider whether a life without a particular capability is a life of dignity on the Kantian understanding of dignity. What Nussbaum seems to mean is that we are supposed to imagine whether or not a life without a specific capability could still be considered a flourishing, good life. This is consistent with Nussbaum’s claim that dignity is not defined prior to capabilities, but rather that the list of capabilities is constitutive of dignity—a.k.a. flourishing.⁴⁹

Given this equation of dignity with flourishing, we are to imagine a life without a certain capability and to ask ourselves whether such a life would still be a flourishing life—a good life for an individual member of that species. But how good? Nussbaum claims to be sketching a minimal account of justice; the guaranteeing of the core capabilities is a minimum requirement for a society to be just. Yet she is not asking whether a life without one of these capabilities is still a life worth living—that would set the bar far too low to generate the kinds of core capabilities she is in favour of. Now that dignity is not a concept defined in advance of the capabilities, and capable of guiding us when trying to answer the question of whether or not a life without a specific capability is still compatible with dignity, the process for identifying core capabilities is more vague. We are to understand a “flourishing” life to be more than just a life worth living, yet less than a fully satisfying life—unless we think that it is a *minimum* requirement of basic justice that every individual lead a fully satisfying life.

When engaging in this process of imagination and reflection, we cannot simply

48. Though this interpretation is unhelpful when applied to Nussbaum’s claim that all humans are equally dignified. Nussbaum, *Frontiers of justice*, 274.

49. *ibid.*, 161–163, 174.

consider whether or not *we* would consider a certain life flourishing in the appropriate sense, for this is meant to be a morally evaluative project. Thus Nussbaum thinks that the capability to harm others—even though it may be typical of a specific species—can never be selected as a core capability because harming others is immoral.⁵⁰ So if we were the kind of people who took particular pleasure in harming others—perhaps we found spiritual delight in the preparation and consumption of uniquely meaty-meals—our judgment that the capability to make and eat meaty meals was essential to a flourishing life would not lead to the inclusion of that capability on the list of core capabilities. To this extent, Nussbaum’s theory is dependent on an antecedent moral theory for guidance. Insofar as Nussbaum is addressing a roughly Rawlsian audience, and trying to build on the notion of an overlapping consensus, this reliance on an antecedent moral theory should not be troubling.

So we have seen that, to determine which capabilities are core capabilities for a given species, we are to imagine a life without a capability and to ask ourselves whether such a life would still count as a flourishing life for a member of that species. If so, then the capability in question is not a core capability; if not, then we must consider whether the capability in question violates an antecedent moral theory. If so, then the capability in question is not a core capability; if not, then it is a core capability. Complicating matters, Nussbaum follows Singer by restricting our concern for capabilities to concern for the capabilities of sentient beings alone.⁵¹

What capabilities are protected? Disabled humans divide into two groups: those who are still leading a recognizably human life and those who are not.⁵² Those in the latter group will be guaranteed different capabilities, likely determined on a case by

50. Nussbaum, *Frontiers of justice*, 366–371.

51. *ibid.*, 361–362.

52. *ibid.*, 181, 187.

case basis. Those in the former group should be guaranteed the same capabilities as all other humans; it is too tempting for society to allow itself to make lesser efforts for disabled humans, and by maintaining a single capabilities list we are able to explain why disabled lives are unfortunate.⁵³ The central human capabilities are roughly:

1. “Life”: To live the typical lifespan of a human.
2. “Bodily Health”: To have good health, adequate nourishment, and adequate shelter.
3. “Bodily Integrity”: To have freedom of movement, to be free from assault, and to have the freedom to procreate.
4. “Senses, Imagination, and Thought”: To imagine and think, to be educated, to express oneself freely, to experience pleasure and avoid pain.
5. “Emotions”: To form emotional attachments to others.
6. “Practical Reason”: To determine one’s conception of the good and plan one’s life.
7. “Affiliation”: To “live with and toward others,” to socialize, to have “the social bases of self-respect,” to not suffer discrimination.
8. “Other Species”: To relate to nature and other species.
9. “Play”: To laugh and play.

53. Nussbaum, *Frontiers of justice*, 179, 182, 188–193.

10. “Control over One’s Environment”: To participate in political choices, to have freedom of speech and association, to hold property, to seek employment, to be free from “unwarranted search and seizure.”⁵⁴

These capabilities are taken to establish both negative and positive obligations—obligations to forbear from acting in certain ways, and to perform specific actions.⁵⁵

Animals would be guaranteed different capabilities based on their species-typical forms of flourishing. While each species would have its own capabilities list, Nussbaum does identify suggestive central capabilities that would be guaranteed to all animals, though they may hold to a greater or lesser extent depending on the species:

1. “Life”: Animals should be free to live, whether or not they take an interest in continued life. Even painless killing counts as a harm, though painless killing should be preferred to painful killing, and euthanasia and culling may be called for.⁵⁶
2. “Bodily Health”: Animals under human control should be protected against cruelty, neglect, inadequate space and nutrition.⁵⁷
3. “Bodily Integrity”: Animals should be free from abuse and harmful treatment (whether or not it is painful). Entitlement to reproductive freedom, with exceptions where violating this freedom is in the interest of the animal.⁵⁸
4. “Senses, Imagination, and Thought”: Animals should be protected from cruel

54. Nussbaum, *Frontiers of justice*, 76–78.

55. *ibid.*, 374–379.

56. *ibid.*, 393–394.

57. *ibid.*, 394.

58. *ibid.*, 395–396.

and abusive treatment, confinement, hunting, and killing; and should be entitled to pleasure, free movement, light and shade, characteristic environments, various opportunities (not just monotony), education and the cultivation of species-typical excellences.⁵⁹

5. “Emotions”: Animals are entitled to form attachments to others, not to be deliberately frightened, not to be subject to solitary confinement or neglect.⁶⁰
6. “Practical Reason”: To the extent that an animal can develop and pursue its own conception of the good, it is entitled to support in this capability.⁶¹
7. “Affiliation”: “Animals are entitled to opportunities to form attachments. . . and to engage in characteristic forms of bonding and interrelationship,” except for violent interrelationships. We have a duty to protect animals from some harms, whether inflicted by humans or by other animals. They are entitled to respect and freedom from humiliation.⁶²
8. “Other Species”: Animals are entitled to relate to nature and other species.⁶³
9. “Play”: Animals are entitled to play, and to the resources necessary for play, like open space and the presence of other species.⁶⁴
10. “Control over One’s Environment”: Animals are entitled to have their interests represented and considered in political decision-making, to respect for habitat, and to respectful work conditions.⁶⁵

59. Nussbaum, *Frontiers of justice*, 396–397.

60. *ibid.*, 397–398.

61. *ibid.*, 398.

62. *ibid.*, 398–399.

63. *ibid.*, 399–400.

64. *ibid.*, 400.

65. *ibid.*

Thus, the capabilities approach generates positive obligations toward animals, and goes beyond concern for pain and pleasure, protecting animal interests that do not involve pain or pleasure.

How does this story about how the core capabilities are selected fit with Nussbaum's account of the foundation or motivation for the capability approach? Nussbaum offers two incompatible accounts of the foundation of the capabilities approach: the first claims that the approach is grounded in a sense of wonder and a love of justice, while the second claims that the approach is grounded in an overlapping consensus.

The “*wonder* interpretation” of the foundations of the capability approach seems to fit best with its method for identifying the core capabilities.

Its [the capability approach's] basic moral intuition concerns the dignity of a form of life that possesses both abilities and deep needs. Its basic goal is to address the need for a rich plurality of life activities. With Aristotle and Marx, the approach has insisted that there are waste and tragedy when a living creature with the innate or ‘basic’ capability for some functions that are evaluated as important and good never gets the opportunity to perform those functions.”⁶⁶

She claims further:

The [capability] approach is animated by the Aristotelian sense that there is something wonderful and worthy of awe in any complex natural organism—and so it is all ready [sic.], in that spirit, to accord respect to animals and recognize their dignity.⁶⁷

66. Nussbaum, *Frontiers of justice*, 346-347.

67. *ibid.*, 93-94.

Finally, while commenting on Aristotle's failure or refusal to draw ethical claims from this sense of wonder, Nussbaum claims:

And yet, if we feel wonder looking at a complex organism, that wonder at least suggests the idea that it is good for that being to persist and flourish as the kind of thing it is. This idea is at least closely related to an ethical judgment that it is wrong when the flourishing of a creature is blocked by the harmful agency of another. That more complex idea lies at the heart of the capabilities approach.⁶⁸

This interpretation appears to be incompatible with Nussbaum's restriction of the sphere of beings whose flourishing deserves moral consideration to sentient beings: if wonder is the source of our moral obligations and we find non-sentient beings wonderful, then it is unclear why our wonder would only ground obligations toward sentient beings. If the starting point is the feeling of wonder and awe, we must recognize that we may feel greater wonder and awe at certain insects and plants than we do at the most primitive sentient beings.

The "*overlapping consensus* interpretation" which Nussbaum appeals to is the idea that the capabilities account could be the object of an overlapping consensus.

The capabilities approach is articulated in terms of a Rawlsian idea of *political liberalism*: That is, the account of entitlements is envisaged as a partial account of the good, for political purposes, which citizens may attach to different comprehensive conceptions of the good. It is articulated, or at least we hope so, in terms of freestanding ethical ideas only, without reliance on metaphysical and epistemological doctrines (such as those of

68. Nussbaum, *Frontiers of justice*, 349.

the soul...) that would divide citizens along lines of religion or comprehensive ethical doctrine. It is therefore hoped that this conception can be the object of an *overlapping consensus* among citizens who otherwise have different comprehensive views.⁶⁹

Yet it is unclear how the wonder interpretation can be presented as the object of an overlapping consensus free of divisive metaphysical and epistemological doctrines. It seems like Nussbaum must either claim that certain things are wonderful, or else claim that humans find certain things wonderful. In the first case, there is the risk that appeal to contentious metaphysical values and epistemological views might be needed. In the second case, it seems that Nussbaum might need to appeal to an aesthetic and moral conception that she must presume all people share. Thus she would be appealing to a specific comprehensive doctrine—one that most people do not seem to share. Thus it appears that there may be some tension between the wonder interpretation and the overlapping consensus interpretation. Moreover, when we examine the method by which capabilities were selected for the list of core capabilities, the method did not seem to involve consideration of what could be affirmed by all comprehensive views, but rather, what seemed to Nussbaum to be wonderful. So, in the face of this tension, I think that the wonder interpretation is the best fit with Nussbaum's arguments, though not a perfect fit.

Thus, Nussbaum's capabilities theory extends considerable negative and positive rights to animals and marginal cases. The "sentience," "wonder," and "overlapping consensus" interpretations guarantee these rights for all sentient beings, all sources of wonder, and for all those on whose membership an overlapping consensus can be

69. Nussbaum, *Frontiers of justice*, 163.

achieved, respectively. These rights are grounded in an Aristotelian notion of dignity or flourishing, and guaranteeing these rights is required for a society to be minimally just.

4.4 Conclusion

This chapter has examined the details of utilitarian, Kantian, and capability theories of rights for animals and marginal cases. Utilitarianism extends the right to equal consideration to all sentient beings, and imposes a duty to maximize aggregate utility on all agents. While we are obligated to promote the welfare of animals and marginal cases equally with human welfare, there are no general proscriptions against the killing or harming of individual animals or marginal cases. Regan's Kantian theory, on the other hand, extends negative rights against harm and disrespect to all subjects of a life. Because harm is not reducible to pain, Regan is able to proscribe painless harms, unlike utilitarianism. While animals hold negative rights on Regan's view, they do not hold positive rights. Finally, capability theory generates positive and negative rights for animals and marginal cases.

With the details of contractarian, utilitarian, Kantian, and capability theories of rights for animals and marginal cases established, we can now go on to evaluate their relative strengths and weaknesses. Chapter 5 will compare the theories from a security perspective, and argue that contractarianism provides the strongest theory of rights for animals and marginal cases.

Chapter 5

Evaluating Competing Theories of Rights

This thesis has now presented the contractarian account of rights for animals and marginal cases, as well as the competing accounts grounded in utilitarianism, neo-Kantianism, and capabilities theory. The current chapter will articulate a set of desiderata to determine which of these accounts offers the strongest theory of animal rights.

5.1 Desiderata For Theories of Rights

5.1.1 Security: The Central Desideratum

A theory of moral rights must have four elements: a set of duty-holders, an argument, a set of right-holders, and a set of rights.¹ We can evaluate a theory of rights by each

1. Susan Dimock pointed out at the *Hobbes and Spinoza in the 21st century: Method, Morality, and Politics* conference that this specification of the elements leaves out one important element to contractarian theories: the circumstances to which the theory applies. I have omitted circumstances as an element, but I have not overlooked them. Circumstances have been omitted because only contractarians discuss them explicitly, making it difficult to compare contractarianism to competing theories in this regard. Circumstances have not been overlooked, however, as I have incorporated them explicitly when considering the contractarian account of what it takes to be a right-holder, the stability of one's membership in the set of right-holders, and the insurance offered by protectorate status. So, while I have not singled circumstances out as an element of theories of rights, I have not overlooked the importance of circumstances to contractarian theories of rights.

of these elements:

1. **Duty-Holders:** Who are the duty-holders specified by the theory? Is the set of duty-holders sufficiently broad?
2. **Argument:** What argument does the theory offer to support its claims? How strong and persuasive is the argument?
3. **Right-Holders:** Who are the right-holders specified by the theory? Is the set of right-holders sufficiently broad and stable?
4. **Rights:** What rights are guaranteed by the theory? Do they protect our most important interests? Do they generate overly onerous duties?

Thus there are at least four ways to claim that an account of rights is impoverished: by claiming that the set of right-holders, the set of duty-holders, the set of rights, or the argument are impoverished. Section 5.1.3 will examine each of these elements in greater detail, and section 5.2 will evaluate how well utilitarianism, Kantianism, capabilities theory, and contractarianism fare on each of these elements.

We might also want to critique a theory as insufficiently rich if it gave us a “flat” rights landscape where there was only one set of rights to hold for one set of right-holders against one set of duty-holders. This would fail to reflect the variety of relations between duty-bearers and right-holders. For instance, family members, co-citizens, or co-nationals may stand in special relations to one another that ground further sets of rights and duties.² As stated in section 1.1, this thesis focuses on basic

2. Donaldson and Kymlicka’s *Zoopolis*, for instance, criticizes animal rights theorists for producing flat theories, and develops a richer account of political animal rights. Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*.

rights which individuals have based on their abilities to help and harm one another, and does not mean to preclude other special rights.

In light of what considerations should we evaluate each element of a theory of rights? To evaluate an account of rights, it is important to begin by asking what it is that we are trying to achieve by articulating these accounts—what are the desiderata of an account of moral rights? I contend that the primary function of an account of rights is a particular form of security, or at least, this is the function that I am primarily interested in. Rights afford us security by convincing others not to interfere in our pursuit of our conception of the good, with particular focus on our most important goals like the preservation of our own lives. On this understanding, rights secure our interests—viz., our desires generally, but with particular focus on our most important desires—against those who would violate them. The primary goal, then, is to secure our interests by constraining the behaviour of others. Rights constrain the behaviour of others by imposing duties on them.

How are these duties *imposed*? Because I am interested in basic *moral* rights, duties must be *justified* to duty-holders. If duty-holders do not already recognize their duties by light of their own reason, we must appeal to reasoned argumentation to persuade them; as a last resort, we may be forced to exclude them from the convention and to coerce them to comply. But compromise and reasoned convention, when available, are almost always to be preferred to exclusion and coercion.³

3. Here I am in complete agreement with Buchanan's conclusion, though I think for different reasons. Where Buchanan starts with an "individualist" ideal, which aims toward anarchy as the best possible order, I do not start from such value-laden premises (See James M. Buchanan, *The Limits of Liberty: Between Anarchy and Leviathan*, vol. 7 (Indianapolis: Liberty Fund, 2000), 5–8.). Instead, I start with the premise that all individuals want to secure their interests as far as possible, and argue that convention offers the best tool to that end. Because of the heavy costs of political solutions, which I discussed in section 3.1.2, it will almost always be preferable not to exclude and coerce people.

I will thus be interpreting philosophical theories of moral rights as aiming at securing compliance with purported duties by offering reasonable arguments that purported duty-holders ought to accept.⁴ This is not a neutral interpretation of the function of theories of moral rights. Some might object that this account of the function of moral theories is already partisan, building in contractarian assumptions that exclude other theories. The point of theories of moral rights is not to be effective, one might object, but to capture objective and irreducible moral truths, or to articulate moral principles that are fair or impartial in some relevant sense of these terms. This risks bringing us face to face with the meta-ethical questions that I am trying to avoid.

To respond without delving into the metaethical, I would claim that even philosophers who endorse other accounts of the function of moral rights must hope that people abide by the moral principles they identify as true, fair, or impartial. I make this claim both as an intuitive argument about what those philosophers are likely to believe, and also more practically as an argument based on the kind of instrumental reasoning at the heart of contractarianism. Those philosophers presumably want to secure the peaceful cooperation of others for the sake of extending and enriching their lives. Such philosophers may then have two separate concerns: one to find a true, reasonable, impartial, or fair account of the principles of justice; the other to actually

4. It is important to note that I do not mean to claim that all ethical considerations aim only at security. The field of ethics is immensely broad, only a subset of which comprises what Gaus calls “social morality.” (Gaus, *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*, 2–14.) Most questions of private ethics may have little to do with security, and more to do with how to live well. Moral rights, I contend, fall squarely within the realm of social ethics. Their distinctive feature is that they aim to secure the compliance of all other agents. I thus take rights to be primarily concerned with security. We can then go further to address such questions as “What would lead to my flourishing?” and “What would be a ‘good’ thing to do?” Such questions will have little to do with security, and do not primarily aim at getting everyone else to act the way we want them to.

secure the peaceful cooperation of their fellows. Addressing one concern may not address the other. Insofar as other moral philosophers are addressing themselves to the task of finding the ‘true,’ fair, or impartial principles of justice, they and I are talking past one another. I am addressing the second question—indeed, I believe the former task to be mostly nonsense, but I cannot argue that here.⁵

Is it fair to evaluate theories of justice that may be aiming at finding ‘true’ moral principles by how well they do at the second task of securing compliance on the ground? I think so. I think those philosophers should not be offended, at least, that those of us trying to address the second question believe that the best place to start is by examining and evaluating theories that attempt to answer the first question. So long as, when evaluating how well moral theories do at the second question, our conclusions are sufficiently qualified and restricted to the domain of the second question, such an evaluation will be fair. I may evaluate how well a theory that aims at ‘true’ morality does at securing compliance, and conclude that such a theory is less effective in this regard; so long as I do not then mistakenly infer that this theory fails to be ‘true,’ my evaluation is fair. Insofar as a philosopher is aiming at irreducible moral truth alone, such a philosopher should not be at all concerned by evaluations of the practical effectiveness of her or his theory. If such a philosopher was concerned, it would perhaps be evidence that the philosopher did indeed care about effectiveness, at which point the evaluation should strike the philosopher as

5. I do not mean to imply that the two concerns are mutually exclusive either. One might argue that ‘true,’ impartial, or fair theories will be the most effective at securing compliance because of their ‘truth,’ impartiality, or fairness. Alternatively, one might pursue multiple independent desiderata with the same theory: one might think that fairness and security are both desirable ends, but that neither accounts for the whole story. Thus, one might be willing to make trade-offs between the two desiderata without being willing to write one or the other off as irrelevant. In this thesis, I am only interested in security. Insofar as others are not at all interested in security, we are speaking past one another. Insofar as others are interested in security as well as other desiderata, my argument may be relevant to them even though they might see it as missing another part of the whole story.

worrisome though not inappropriate. If a philosopher is indeed concerned about effectiveness, then an evaluation based on effectiveness is fair game. If a philosopher is not concerned about effectiveness, then those of us who are so concerned should be free to talk amongst ourselves without bothering the philosopher in question, even if we are talking about her or his theory.

What about circularity? By interpreting moral theories as aiming at security, am I not appealing to the very premises from the very arguments that I am trying to justify? The human desire for security was a fundamental premise in the mutual advantage argument for rights. By using the premise now to evaluate competing theories, or even just to further bolster the mutual advantage theory, I am begging the question. This objection may be particularly glaring because it is precisely the kind of objection that I would like to level against my opponents: to build a theory of justice as impartiality, and then to use impartiality to evaluate competing theories, would seem to be viciously circular. Why should the same not be the case for my argument?

I do not find the circularity objection pressing in the mutual advantage case because the premise in question is meant to be a descriptive empirical claim. In the contractarian account of rights, the premise functions as an empirical description of humans and what, given their nature, *moves* them to act in various ways. Knowledge of what drives humans allows us to identify principles that each individual should (predictively) follow. It is thus unsurprising that, when trying to evaluate competing theories *from a human perspective, based on what we want from an account of rights*, the same empirical premise should appear. Thus, while this may be circular, the circularity is not vicious, and indeed is inescapable insofar as the empirical premise in

question is true. What else could I appeal to, given my denial of irreducible natural moral facts? Only empirical claims about humans and our desires. However, while I did aim to articulate an independent approach to evaluating competing theories of moral rights, I must admit that this may merely amount to a rephrasing of the contractarian argument designed to appeal to a broader audience. Should this be true, I hope that this way of representing contract theory makes the differences between contractarian and non-contractarian theories more explicit, and allows for a better comparison of their relative strengths and weaknesses.

The rest of this section will proceed to evaluate the different features of competing accounts of rights based on the goal of achieving compliance by reasoned argument. Again, not all moral theories need endorse reasoned compliance as a desideratum, and evaluating these theories based on different desiderata may lead to different conclusions concerning the relative strengths of each theory.

5.1.2 How Do “Protections” Fit In?

In sections 3.1.6 and 3.1.7 I introduced an indirect form of moral consideration that I called “protectorate status.” How do protections fit in to a theory of rights? Protections provide the same kind of security as rights: both provide security by imposing duties on others. Thus, when evaluating how well a theory does in terms of providing security, I think that it would be appropriate to consider both rights and protections.

5.1.3 Criteria For Evaluating Elements of Theories of Rights

This subsection will apply the general desideratum of security to each of the elements of a theory of rights to derive more specific criteria for evaluating the strength of

competing theories with respect to each element. With these criteria determined, section 5.2 will evaluate the competing accounts of rights considered in this thesis, and argue that contractarianism is the strongest.

The four elements are structured by the claim that the principal desideratum of a theory of rights is security achieved through reasoned argumentation. This claim specifies the audience that the argument ought to address: the argument for a theory of rights ought to address the set of duty-holders, as it is by persuading the duty-holders to recognize rights that a theory of rights achieves its primary aim. Thus the argument aims to justify the set of rights and the set of right-holders to the desired set of duty-holders. The first element to examine, then, should be the set of duty-holders. If the other elements must be justifiable to the desired duty-holders, we must know who the desired duty-holders are in order to know what they will find justified.⁶ This section will apply the central desideratum to each of the elements to derive criteria specific to each element.

Duty-Holders

I take one of the key features of moral rights to be the desire that the set of corresponding duty-holders should be as comprehensive as it can reasonably be. We can address the question of what actions would be “morally good” (whether or not they are morally required), for instance, without consideration of how many people

6. Note that this structuring does not apply to all theories of rights; it arises from the central desideratum of security. If, for instance, we were articulating a theory of rights intending to capture and adequately reflect objective moral truths, then this structuring would not apply. It would be perfectly acceptable, in such a case, to start with assumptions, considered judgments, or moral premises that were not shared by all of the purported duty-holders so long as those premises were *true* in the relevant sense. And the argument in that case need not justify anything to the relevant set of duty-holders—it can be targeted only at those other theorists who are already sufficiently advanced in their pursuit of moral truth.

will agree, or will perform those actions based on our arguments. Likewise, we can address the question of what human (or butterfly, or dog, etc.) flourishing consists in without addressing duties or compliance at all.⁷ But talking about moral rights is fundamentally different: the aim of the game when articulating a theory of moral rights is to get as many agents as possible to recognize and respect those rights. Talk of moral rights is tightly bound to talk of duties applying to a particularly expansive set of duty-bearers. If someone advanced a theory of basic moral rights, but claimed not to be talking about duties at all, or to be talking about duties but duties that only applied to a handful of people, many would find the use of term “rights” strained and worry that this person was not truly talking about rights at all.⁸

Setting aside concerns about our common ways of using the term “rights” and intuitions about them, I can also explain why we want as expansive a set of duty-holders as possible based on my account of the function a theory of moral rights.⁹ Rights provide us with a form of security by imposing duties on others, and that is why we want rights to begin with. The greater the set of agents who have duties to respect our rights, the greater the security with which those rights provide us. So we want rights to generate duties that apply to the greatest possible set of individuals. Thus, if one account of rights offers a narrower set of duty-holders than another account, the former is stronger than the latter with respect to the set of duty-holders.

7. I do not mean to suggest that all philosophers who talk about flourishing do not thereby intend to talk about duties and compliance. I mean only to suggest that we could intelligibly talk about flourishing that way, whether or not some people want to appeal to flourishing as grounds for rights and duties.

8. This is certainly more controversial. It makes sense, after all, to say that the contract between A and B gave B a right but no one else. But this thesis is concerned with *basic* moral rights. See section 1.1 for the sense of “basic.”

9. After all, we are not simply impartial theorists who stand outside the system—we are candidate right-holders ourselves, who want to benefit from the system in the same way that all of the other members of the system do.

Argument

Given that we want the set of duty-holders recognizing our rights to be as broad as possible, that we want the duties of duty-holders to be particularly strong, and that the way to get what we want is through reasoned argument, we need an argument that will be as persuasive as possible to the broadest set of potential duty-holders. This amounts to two criteria: of two competing theories, the one that is persuasive to the widest set of duty-holders is, *ceteris paribus*, the one with the stronger argument. And of two competing theories, the one that is most persuasive (*viz.* that gives the strongest reasons—reasons that are not easily outweighed by competing reasons), *ceteris paribus*, has the stronger argument. Because we want the argument to be persuasive to the broadest possible set of duty-holders, knowing that duty-holders endorse divergent conceptions of the good, an argument should be persuasive to as many different conceptions of the good as possible.

Right-Holders and Protection-Holders

How should we evaluate theories based on their accounts of the right-holders? Given the desire for security, we want the set of right-holders to include ourselves and all those whose security we desire. Moreover, we want our membership (and that of those whose security we desire) to be as permanent as possible. This gives us criteria for determining when the set of right-holders is sufficiently broad, and when membership in the set of right-holders is sufficiently stable (when it is unlikely that we or those whose security we desire will lose our membership).¹⁰

10. Donaldson and Kymlicka have raised concerns about stability against theories that ground rights in personhood. See Donaldson and Kymlicka, *Zoopolis: A Political Theory of Animal Rights*, 27; Kymlicka, *Contemporary Political Philosophy: An Introduction*, 134–136.

Our desire for security may also provide us with one criterion by which to assess the upper limits of the set of right-holders. Insofar as our desire for security is threatened by the expansion of the set of right-holders, we will reasonably resist the inclusion of what I will call “threatening right-holders” in the set of right-holders. For example, should someone propose adding all of our food sources to the set of right-holders, and affording them rights against our eating them, this would violate our desire for security. I do not mean to be attributing this position to any philosophers; however, there may be more subtle ways in which this kind of trade-off might arise. For instance, should we face policy decisions where fulfilling purported obligations toward animals or ecosystems would lead to worse outcomes for humans, either by direct consequence of the policy adopted (for example, if by limiting human development in natural habitats, we perpetuated or aggravated a housing affordability and availability crisis) or due to foregone opportunities (for example, if the money we spent on the initiative in question would otherwise have gone to saving human lives), this would equally threaten our security. It is important to note that this limit does not just arise when considering whether to add animals or ecosystems to the set of right-holders. In any case where adding anyone—human or non-human—to the set of right-holders threatened our security, we should resist such an addition.¹¹ Thus, we want all threatening right-holders excluded from the set of right-holders.

When evaluating a theory based on its set of right-holders, we should also include in our evaluation the set of protection-holders. Because protections rely on the same logic as rights, their justification is equally strong and there is no need to consider protectorates separately. Thus, we can combine the set of right-holders and

11. This is connected to the idea of the circumstances of justice. Should we find ourselves in a situation where any constraint on our part would be against our interest, justice-based constraints would have no place.

protection-holders into one set to be evaluated together.

Thus, we have the following 3 criteria for evaluating competing theories of rights with respect to their sets of right- and protection-holders:

1. *Ceteris paribus*, of two competing theories, the one that includes more of the people whose safety we desire (including ourselves) in its set of right-holders or protection-holders is strongest with respect to right- and protection-holders.
2. *Ceteris paribus*, of two competing theories, the one that makes the membership of those whose safety we desire most stable (i.e., the least easy to lose) in the sets of right-holders and protection-holders is strongest with respect to right- and protection-holders.
3. *Ceteris paribus*, of two competing theories, the one that includes the fewest threatening right-holders in its sets of right-holders and protection-holders is strongest with respect to right- and protection-holders.

The first criterion tells us when the set of right- and protection-holders is too narrow, the second tells when membership is insufficiently stable, and the third tells us when the set is too broad. All three are “*ceteris paribus*” criteria because they all claim to evaluate theories with respect to the same element. A complete account would tell us the relative importance of each criterion.

Rights

What interests should rights secure? Here we are pulled in two directions. On the one hand, we want as many of our interests as possible to be secured. But on the other hand, given the awareness that we will also be duty-bearers, we want to ensure

that the duties we face are not excessively onerous. As other people are capable of constraining our ability to pursue our interests by their behaviour, so too may they constrain our pursuit of our interests by gaining rights against us. These two ways in which people can threaten our security are the same in kind, and thus both must be considered when evaluating theories of rights from a security perspective.

But this last point—that we want duties not to be overly onerous—is easily overstated for the following reason: recognizing the rights of others can make our lives go better, even setting aside the benefits we reap directly from holding those rights against others. Recognizing some rights for others can contribute to social stability, allowing for all sorts of social cooperation that benefits us. As Hobbes noted, the state of nature is not only one of constant fighting, but also one deprived of industry, science, and culture (I.xiii.9). Even if it were possible to arrange to be the only one who held rights, this would not be desirable—even from one’s own perspective—because it would preclude the kind of social cooperation that enriches the lives of all. We may still be pulled in both directions with respect to some rights, but with respect to the rights needed for social cooperation to be possible, all should want those rights to protect everyone even knowing that they will be duty-bound by these rights. So when evaluating theories of rights, we should prefer that theory which achieves the best balance between our desire that our rights should be generous, and our desire that duties should not be overly onerous.

5.2 Evaluating Competing Theories

Given the above criteria, how do the relevant competing theories compare to one another? This is a question of balancing the strengths and weaknesses of each theory,

and the examination here produces the kind of mixed results typical of such balancing acts. I contend that contractarianism does well with respect to the desiderata for each element. Competing theories do better than contractarianism with respect to the stability of membership in the set of right-holders. They provide more generous sets of right-holders, though it is not clear that they do better on the sufficiency desideratum as this generosity may be more than sufficient. And they provide more generous sets of rights, making them better from the perspective of right-holders, though they do this at the expense of the perspective of duty-holders and thus may not achieve the best *balance* between the two roles we occupy. Contractarianism, however, does best from the perspective of duty-holders and achieves the best balance between the two perspectives in the set of rights it generates. Contractarianism is also the least likely theory to contain threatening right-holders. But most importantly, contractarianism truly stands out as addressing the broadest set of duty-holders, and as giving the strongest reasons for duty-holders to fulfill their obligations. I acknowledge, in section 5.2.5, that this argument is importantly incomplete, and as such, can only be suggestive, not deductively valid; however, I do think that the argument offers strong reason for concluding that contractarianism is the best theory from the security perspective.

The limits of this argument must once again be emphasized: I contend that contractarianism is best from a security perspective, and thus offers the best account of rights and duties. It does not provide a complete picture of morality: other theories may provide better accounts of moral goodness and badness and of how to lead a good, flourishing life. I think the best solution is to allow each theory to perform the task at which it is best: contractarianism offers the best account of rights, while other

theories offer better accounts of moral goodness and flourishing, etc. One should be free to act according to accounts of moral goodness and flourishing so long as one does not violate the rights of others as determined by contractarianism.

Because I am primarily interested in defending contractarianism, I will focus on comparing other theories' performance to that of contractarianism. Implications for how the alternative theories fare relative to one another will only be discussed when it is relevant to the evaluation of contractarianism. The rest of this section will proceed to examine each theory in turn, considering how well it does according to each of the element-specific desiderata.

5.2.1 Utilitarianism

Duty-Holders and Arguments

Because rights guarantee our security by persuading duty-holders with rational argument that they ought to fulfill certain duties, the strength of a theory's argument will determine the set of duty-holders that the theory can address. Thus, the strength of the competing arguments and their sets of duty-holders must be examined together. Because we want security against the widest set of duty-holders possible, we want an argument that will speak to the widest audience possible, and that will give them the strongest reasons to restrain themselves.

Utilitarianism appeals to three fundamental principles: the sentience principle, the basic principle of equality, and the duty to maximize expected aggregate utility.¹² The sentience principle claims that if a being is sentient then it has interests, namely,

12. Singer, *Animal Liberation*, 2–9, 17.

interests in avoiding pain and pursuing pleasure.¹³ These interests are morally important.¹⁴

The second fundamental premise, the basic principle of equality, claims that all interests ought to receive equal consideration. To claim that animal suffering is less morally important than human suffering, even where that suffering is acknowledged to be of the same quality and quantity, or to claim that your suffering is less morally important than mine, would violate the basic principle of equality.

The final premise claims that we should act so as to maximize the expected net pleasure and minimize the expected net suffering of all sentient beings.

A great deal of the job of arguing for moral consideration for animals is done by the first three premises—all that remains to be done is to add an empirical premise that many or all animals are sentient. But should we accept these three premises?

Firstly, we should note that Singer takes the task of justifying moral claims in the face of skepticism seriously, carrying on a utilitarian tradition that goes back at least as far as John Stuart Mill. Mill claimed that one must identify the “sanction” of a moral theory—the reason for thinking that theory is correct, and that one should obey its dictates—whenever one is calling on someone to adopt a moral theory that he or she was not raised to find natural.¹⁵ Singer explicitly agrees that received ethical ‘wisdom’ and ethical intuitions can and ought to be questioned, and that accounting for the

13. Utilitarians may want to make the further claim that sentience is necessary for a being to have morally relevant interests. For the purposes of this argument, we need only consider the sufficiency claim.

14. Utilitarians may want to make the stronger claim that all *and only* interests in avoiding pain and pursuing pleasure are morally relevant, denying the relevance of any ‘interests’ that are not grounded in considerations of pleasure and pain. But for the purposes of this argument, we need only consider the weaker claim that *all* interests in avoiding pain and pursuing pleasure are morally relevant.

15. Mill, “Utilitarianism,” 203

foundations of ethical systems is important.¹⁶ Indeed, Singer raises the concern that neo-Kantian ethics may become a closed system, refusing to question its foundational premises, and incapable of justifying them.¹⁷ So what argument does Singer offer to ground the three principles in question?

As noted in section 4.1, Singer does not take special pains to defend the principles in *Animal Liberation*, but presents an argument in *Practical Ethics*. He argues that ethics requires us, at a very basic level, to justify our behaviour to others, and that such justification requires us to take up a universal perspective where we give everyone's interests equal consideration.¹⁸ Given that each individual is interested in maximizing his or her own net utility—that is, in seeking out as much pleasure as possible while avoiding as much pain as possible—justifying our actions to each perspective will require us to consider the interests of all sentient beings, and to show that aggregate utility was maximized.¹⁹

This is arguably the standard defence of utilitarianism. The argument can be found in Mill, who claims that each human values her or his own happiness, that any other interests individuals have are instrumental to the interest in happiness, and thus that “the general happiness, therefore, [is] a good to the aggregate of all persons.”²⁰ Smith, Hume, Sidgwick, and Harsanyi have taken it to be an essential or definitional element of morality that it is tightly bound up with impartiality, and Mill's “aggregate of all persons” can be equated with Sidgwick's “point of view of . . . the universe,”²¹ and Hume's “judicious spectator.”²² We need to take up these perspectives because,

16. Singer, *Practical Ethics*, 4–8.

17. *ibid.*, 285–286.

18. *ibid.*, 9–10.

19. *ibid.*, 11–12.

20. Mill, “Utilitarianism,” 210.

21. See book III, chapter xiii, section 3. Sidgwick, *The Methods of Ethics*, 382.

22. See III.iii.1. Hume, *A Treatise of Human Nature*, 632.

were we each to evaluate things from our own perspectives, our judgments would disagree. My interests may lead me to judge differently than your interests lead you to judge. Insofar as we seek agreement in our moral judgments, we need to take up a shared perspective, and the only shared perspective that each of us could agree to is one wherein each of our interests are taken into equal account.²³

But there are three premises employed in this defence of utilitarianism that are deeply questionable. The first is the assumption that behaving ethically requires us to justify, or seek to justify, our actions to others. This is virtually the reverse of the Hobbesian contractarian starting point: whereas Hobbes denied the existence of natural moral or political authority, the utilitarian starting point seems to grant natural moral authority to *everyone*. Thus, contractarians start with a system of unlimited natural right,²⁴ and justify *constraint* to parties in such a state. Utilitarians, on the other hand, start with a system of *no* right, and insist that all actions be justified to all others. The question must thus be raised: Why should we accept the assumption that we ought to justify all of our actions to everyone? Here, I think, the options for defence are limited: it must either be labeled a fundamental intuition or be claimed as a shared ethical judgment or practice, but both options run afoul of Singer's own commitment to questioning moral intuitions and received moral wisdom and to justifying ethical claims.²⁵ Indeed, I think that the claim that we must justify our actions to others is already a moral claim. Moreover, insofar as we consider the utilitarian characterization in the context of social morality evaluated from a security perspective, we can see that it is backward. We stand to benefit from an ethical

23. Hume, *A Treatise of Human Nature*, 625-652 (III.iii.1-2).

24. Bear in mind, "right" is here being used in a non-moral sense: having an unlimited right to all things amounts to not having any obligation to forebear from anything. See section 2.1.4.

25. Singer, *Practical Ethics*, 285-286.

system primarily insofar as we are right-holders; it is primarily duties that need to be justified to duty-holders by right-holders, as I argued in section 5.1.3 on page 168.

The second problematic premise is the claim that, in order to justify our interests to others, we must adopt and reason from a universal or impartial perspective. This claim stands out when Hume makes it, in particular, because he himself had already shown it to be false: insofar as Hume successfully showed that individuals could be brought to agree to a convention, each from a self-interested perspective, he showed that a universal or impartial perspective was not needed to achieve consent in the moral evaluation of actions or rules.²⁶ Insofar as mutual advantage contractarianism is successful in showing that individuals should agree to rights and duties without appealing to a universal perspective, the second premise is false.

The final problematic premise is the claim that justifying action from a universal or impartial perspective requires us to maximize *aggregate* utility, or even to focus on aggregates more generally, whether or not we are maximizing. Singer himself is keenly aware of the problem of appealing to universality: either the notion of universality will be so thin that it includes incompatible ethical theories, or else it will be so thick that it assumes a particular ethical theory without warrant.²⁷ This problem is further aggravated by the fact that Gauthier himself offers a Humean argument that the impartial perspective would endorse Gauthier's mutual advantage contract theory.²⁸ The primary point of contention between utilitarian impartiality theories and Gauthier's is on the issue of aggregation: would an impartial spectator allow losses to some for the sake of gains to others, or would an impartial spectator insist

26. Hume, *A Treatise of Human Nature*, 625-652 (III.iii.1-2).

27. Singer, *Practical Ethics*, 11.

28. Gauthier, *Morals by Agreement*, 233-267.

on mutual benefit?²⁹ Gauthier and Rawls both argue against the utilitarian view on this issue, though for different reasons.³⁰ Gauthier claims that each individual must be able to identify with the impartial decision made, but no individual has a reason to continue to participate in social cooperation if it poses a loss to the individual compared with non-cooperation.³¹ Thus, he contends that impartiality leads to mutual advantage.

To bolster the argument that we must justify our actions from a universal perspective by showing that they maximize utility, Singer argues that we can justify adopting an ethical perspective from a non-ethical perspective, after which we need no longer consider the non-ethical perspective.³² Singer argues that from the non-ethical perspective we should be able to recognize that the life of a psychopath—“a person who is asocial, impulsive, egocentric, unemotional, lacking in feelings of remorse or shame or guilt, and apparently unable to form deep and enduring personal relationships”³³—is a meaningless, unfulfilling life, whereas the ethical life is meaningful and fulfilling. Thus, from the non-ethical perspective, we can justify taking up the ethical perspective, which entails accepting at least the first two questionable assumptions discussed immediately above, if not all three. Singer claims that the psychopath’s life is meaningless because:

Most of us would not be able to find full satisfaction by deliberately setting out to enjoy ourselves without caring about anyone or anything else. The pleasures we obtained in that way would seem empty and soon pall. We

29. Gauthier, *Morals by Agreement*, 240, 244–261.

30. *ibid.*, 240–245; Rawls, *A Theory of Justice*, 23–24, 163–166.

31. Gauthier, *Morals by Agreement*, 244, 248–249, 254, 257–259, 261.

32. Singer, *Practical Ethics*, 268.

33. *ibid.*, 288.

seek a meaning for our lives beyond our own pleasures and find fulfillment and happiness in doing what we see to be meaningful.³⁴

Moreover, Singer claims that we would not avoid the meaninglessness of the psychopath's life by adopting long-term but egoistic goals—goals rooted in a desire for our own happiness. Even adopting such goals, we would face the following dilemma: either we would satisfy all of those goals, in which case we would grow bored and unhappy, or we would continue to replace fulfilled goals with new goals, thereby becoming insatiable.³⁵ Thus, we would still be trapped in a meaningless and unsatisfying life; the only way out is to adopt the ethical point of view.

There are five main flaws with this argument. Firstly, the kind of mutual contract theory developed here is not addressed exclusively to psychopaths: of course people care for *some* other people, and of course they can have emotions and feelings of guilt. Contractarianism claims only that our concern for others is limited; it neither applies to all individuals equally (or at all), nor is it always strong enough to motivate us to sacrifice our own interests.³⁶ Moreover, contractarianism relies on people not being overly impulsive—they must be able to take an interest in their long-term well-being, and to act so as to best pursue their long-term interests. And of course, people may feel guilt when they break the moral rules: the question is, in response to violations of which rules should we feel guilt? The contractarian would fill in the rules differently than a utilitarian, but would not reject guilt completely. Thus, insofar as Singer is targeting mutual advantage ethical theories, and not just amoralists, Singer may be guilty of building a straw man by claiming that only psychopaths, so characterized,

34. Singer, *Practical Ethics*, 291.

35. *ibid.*, 292–293.

36. See section 2.2.1 on page 60.

would find such theories relevant. Even insofar as Singer successfully shows that we should not choose to be psychopaths, he has not shown that one should be a utilitarian as opposed to a mutual advantage contractarian. While the psychopathic life may be meaningless, the life of a contractarian is relevantly different from the psychopath, and may not be meaningless.

Secondly, Singer has not shown that it would be meaningless to stick exclusively to egoistical goals. Indeed, a similar view has been defended by Steven Luper, who argues that we should “epicureanize” our desires by mostly only taking on projects that we think we can complete in our lifetime, and taking steps to ensure their completion as we near our end, but by delaying the completion of our projects as long as possible.³⁷ The bare possibilities of insatiability or boredom do not show that egoistic lives must fall into one of those possibilities. Moreover, it is not clear why these possibilities do not apply to non-egoistic lives: non-egoistic interests seem as prone to fulfillment as egoistic interests, so it seems just as possible that a non-egoist could become insatiable. Likewise, if all of the non-egoistic projects that I had undertaken were completed and I did not replace them with new projects, could I not also grow bored? Thus, Singer has not shown that the egoistic life is meaningless in any special sense that does not apply equally to non-egoistic lives.

Thirdly, the “ethical” life that Singer advocates for is not sufficiently connected to the three fundamental assumptions of Singer’s argument for utilitarianism. It is unclear how we get from the ‘ethical’ life that would avoid the pitfalls of psychopathic

37. For the sake of simplicity, I have glossed over some of the details of Luper’s argument that are less relevant to the argument at hand. For instance, Luper thinks that we should be willing to take on some projects that are bound to outlast our lives, for instance having children who have good lives, but that we should take steps to ensure that these projects are fulfilled without us after we die, for instance by equipping our children to take care of themselves. Steven Luper, “Annihilation,” in *Life, Death, & Meaning*, Second, ed. David Benatar (Toronto: Rowman & Littlefield, 2010), 201–222.

and egoistic lives and make our lives meaningful, to the claims that we must justify our behaviour to others from a universal perspective by maximizing aggregate utility. We can imagine a whole host of ways to live—including ways that are neither psychopathic nor utilitarian, but are consistent with the mutual advantage contract theory developed here—that would be meaningful. Trying to connect this argument to utilitarianism, and to exclude contractarianism, is impossible.

Fourthly, it is highly questionable whether the kind of conversion that Singer discusses is possible. If we adopt the non-rational view-point and ask whether it is rational to adopt the ethical viewpoint, we should also consider some alternative viewpoints. Consider the following dilemma: either it will be rational to do everything that the ethical perspective asks one to do, or else there will be some things that the ethical perspective asks one to do that will not be deemed rational from the non-ethical point of view. Under the first scenario, it is not clear what work the ethical perspective is doing; in that case, we seem to have a complete reduction of ethical to rational, which is much more than Singer wants—indeed, that is precisely what Singer is objecting to! Under the latter scenario, we would only choose the ethical perspective if that was our only choice. But why not choose a mixed strategy, deciding to do only those things that the ethical perspective requires which are also deemed rational from the non-ethical perspective, and otherwise sticking with the non-ethical perspective? So we could endorse caring for others directly, but resolve not to do so when doing so would cost us more than we stood to gain. So if we were allowed to choose strategies other than the pure ethical strategy, it would be rational to choose a mixed strategy, which seems indistinguishable from sticking with the non-ethical

perspective.³⁸

Finally, the argument is weak, and Singer acknowledges this. He notes that “All this is speculative. You may accept or reject it to the extent that it agrees with your own observation and introspection.”³⁹ Furthermore, while seeming to discuss both his own views and those of Henry Spira, he admits that the argument

will not provide everyone with overwhelming reasons for acting morally. It cannot be proven that we are all rationally required to reduce pain and suffering and make the world a better place for others. Ethically indefensible behaviour is not always irrational. We will probably always need the sanctions of the law and social pressure to provide additional reasons against serious violations of ethical standards.⁴⁰

Different people find different things meaningful. Based on the four objections presented above, it does not seem that Singer’s argument has given good reason for many people to adopt a utilitarian perspective, let alone ruled out a contractarian perspective, or even a very narrowly egoistical perspective with no care for anyone else. The argument is likely only appealing to those who already endorsed the three premises grounding Singer’s argument for utilitarianism—that we should justify our behaviour to everyone from an impartial perspective by maximizing aggregate utility. This will resonate with some people, but so will Kantian arguments resonate with people who already share their moral intuitions, and Singer has already rejected this

38. Indeed, it is not clear what would distinguish adopting the mixed strategy from adopting the non-ethical strategy, which may be a *reductio* of the whole thought experiment. The argument that I have developed in this paragraph is largely inspired by Hampton’s objection to the “conversion” interpretation of the act of authorizing a sovereign in Hobbes’s *Leviathan*. Her argument could not apply directly to Singer’s argument, so I made some modifications, but my argument is still structurally similar. Hampton, *Hobbes and the Social Contract Tradition*, 208–220.

39. Singer, *Practical Ethics*, 292.

40. *ibid.*, 295.

method of ethical reasoning. Because Singer is not able to offer strong arguments supporting the premises that differentiate utilitarianism from mutual advantage contractarianism, the grounds for being utilitarian as opposed to contractarian must lie in starting assumptions and intuitions. Thus, Singer's argument will only appeal to those who already held the utilitarian assumptions.⁴¹

This is not to claim that Singer's moral views, or the intuitions and beliefs on which they rest, are false; the absence of a good argument in favour of a conclusion does not show the conclusion to be false. But I am not questioning which moral theory is "right," but which theories will best secure us against harm. Until utilitarianism can support its grounding assumptions, it will have no way of securing us against people who do not endorse those assumptions.⁴² Insofar as theories of rights are measured by their abilities to secure us through reasoned argumentation, this exclusion risks making utilitarianism nearly irrelevant: the point is to persuade people who would otherwise be inclined to cause us harm, and such people likely do not endorse the two basic principles. Utilitarianism can only address people who would be disinclined to harm us anyway.

How strong are the reasons that utilitarianism provides for duty-holders to fulfill their duties? This question is quite broad; I will attempt to answer the more manageable question: Does utilitarianism provide stronger reasons for duty-holders to fulfill their duties than contractarianism? The main difference between the two theories is the fundamental principle of equality, which leads utilitarianism to call for agents to make uncompensated sacrifices, whereas contractarianism insists that moral

41. The 'weakness' of Singer's argument should not, however, be exaggerated. The seriousness with which he attempts to ground ethical obligations rivals that of contractarians, and greatly exceeds that of the other competing theories considered here.

42. If no such people existed, this would not be a problem, but it seems that such people do exist.

requirements must be in the interest of the duty-holders. Because of this difference, *akrasia* will be a much larger problem for utilitarianism than for contractarianism. Utilitarianism can require agents to perform actions that are counter to their interests, giving them a strong incentive to fail to fulfill their duty. This situation cannot arise for contractarianism. Contractarianism does face other forms of *akrasia*, like doing what is in your short-term interest but against your long-term interest, but it can never run against one's long-term self-interest. Thus, *akrasia* will be a greater problem for utilitarianism than for contractarianism. In this respect, utilitarianism offers weaker reasons for duty-bearers to fulfill their duties; the reasons utilitarianism offers are more prone to be overridden by weakness of will than the reasons offered by contractarianism. This should trouble us from a security perspective: insofar as we want a theory to secure our interests against others, theories that offer reasons that are more prone to being overridden are less effective at generating security.

Right-Holders

For utilitarianism, the set of right-holders contains all sentient beings.⁴³ The utilitarian account is thus rather straight-forward. The vast majority of animals will be included in the set of right-holders, including marginal cases.

Is the utilitarian set of right-holders sufficiently broad? There is no objective way to answer this question; some people are more caring—*viz.* desire the safety of a greater set of individuals—than others. The criterion is fundamentally subjective. Thus, each individual will answer the question differently. However, we can still make objective predictions about people's subjective evaluations.

It is important to note that this criterion is asking for a fact about one's subjective

43. Singer, *Animal Liberation*, 7–8.

attachments, not about one's moral beliefs. Thus a particular utilitarian may believe that, as a matter of morality, all sentient beings are owed equal moral consideration, but may simply not be subjectively moved in ways that match these moral considerations. The utilitarian may have a human arch-nemesis, and should that nemesis be assaulted by a gang of thugs, the utilitarian may not mind. Indeed, we could rephrase the criterion in a manner that was much more useful for a utilitarian: a theory of rights that protects all those whose wellbeing makes one happy (or satisfies one's desires), and whose suffering makes one unhappy (or violates one's desires), is better than a theory that does not. Thus, I think that there are strong reasons to think that this criterion will not simply reproduce theoretical disagreements. A utilitarian is likely to be satisfied by a less comprehensive set of right-holders than the sentient set generated by utilitarianism.

Because the first criterion asks for a subjective assessment of the set of right-holders, assessments in light of this criterion will not always match theoretical commitments. Insofar as answers do not match theoretical commitments, I contend that it will approve of less comprehensive sets of right-holders. Hume, in his utilitarian thought, and Mill both agreed that individuals need not be—and likely would not be—subjectively moved by the principle of utility.⁴⁴ Because subjective concern extends to a narrower set of individuals, this will have the effect of leading utilitarians to judge sets of right-holders that are less comprehensive than the sentient set to satisfy the first criterion as being sufficiently comprehensive. Thus, there will be a tendency for utilitarians to judge less comprehensive sets of right-holders to be sufficient, and there is the possibility that utilitarians may judge the contractarian

44. Hume, *An Inquiry Concerning the Principles of Morals*, 19, 55; Hume, *A Treatise of Human Nature*, (III.ii.1) 532–536; Mill, “Utilitarianism,” 195.

set of right-holders to be sufficient. Indeed, insofar as the subjective attachments of utilitarians extend only to other humans and to certain domesticated animals, the contractarian set of right-holders is likely to be deemed sufficient.

How stable is one's membership in the utilitarian set of right-holders? Utilitarianism extends moral consideration to all sentient beings, so one's membership is as secure as one's sentience. Membership in the utilitarian set of right-holders, then, is quite stable—indeed, it is the most stable of all the theories under consideration in this thesis.

Does utilitarianism include threatening right-holders—that is, right-holders whose membership threatens the security of humans?⁴⁵ The utilitarian set does not seriously threaten human security so long as all humans can be adequately nourished, clothed, and sheltered without violating the moral duties owed to sentient animals. This condition seems to hold, so the set of sentient beings is not seriously threatening to humans.

Nevertheless, the utilitarian set of right-holders may be seen as containing threatening members, even though the threat is not so severe as to imperil human life. The kinds of sacrifices that Singer views as mandatory according to utilitarianism, whether for the sake of humans or of other sentient beings, would indeed result in a loss of utility for the purported duty-holders (even if they maximized aggregate utility). Thus, duty-holders can reasonably see the prospect of utilitarianism as a threat to their happiness; unless an argument can show that they stand to gain more than they stand to lose, they can see utilitarianism as threatening. It is impossible

45. I say “humans” because it is human behaviour that we are trying to modify by articulating arguments for moral rights, and thus it is their security that matters when evaluating threatening right-holders.

to separate the threat originating in the set of right-holders from the threat originating in the set of rights: should the set of right-holders be held constant but the set of rights reduced, utilitarianism could be less threatening to duty-holders. But the two sets are inextricably linked by the utilitarian argument specifying the two by reference to the moral relevance of pain and pleasure. While I have distinguished the elements in a theory of rights for the sake of evaluation, the elements are tightly connected and in some cases the distinction between elements is rather artificial. To compensate for this, I will need to heavily qualify the conclusion in this case: the utilitarian set of right-holders is not so threatening as to imperil the lives of duty-holders, but taken together, the utilitarian set of right-holders and of rights do call for considerable sacrifices on the part of duty-holders. According to the utilitarian reasoning, these sacrifices are more than compensated for by the gains of others, but the sacrifices made by the duty-holders cannot reasonably be expected to be balanced out by gains *to the duty-holders* reaped from the moral system. Thus, taken together, the utilitarian set of rights and of right-holders are threatening to many duty-holders, though not to the extent of imperiling their lives.

This assessment of utilitarianism as threatening is reinforced when contrasting utilitarian reasoning with contractarian reasoning. Whereas contractarianism insists that duties be shown to be in the interest of duty-holders, so that duty-holders themselves can expect that fulfilling their duties will make their own lives go better, utilitarianism justifies duties by reference to aggregate utility, not to that of the duty-holders. Thus, utilitarianism can call on a duty-holder to sacrifice his or her utility, without the prospect of further utility gains for the duty-holder, for the sake of improvements in some other individual's utility—someone whom the duty-holder

may never meet or, worse still, someone whom the duty-holder may positively despise. Thus the utilitarian set of right-holders and of rights is bound to be more threatening than those of contractarianism.

Rights

This thesis is limited to evaluating competing accounts of rights around the “margins” or at the “frontiers.” To evaluate competing theories in terms of the set of rights they generate would require us to evaluate complete accounts of rights, taking us beyond the scope of this thesis. However, by considering the pattern of reasoning employed by each theory, I can comment on how likely each theory is to be successful in balancing our interests as right-holders with our interests as duty-holders.

From the perspective of right-holders, utilitarianism generates a set of rights that is more generous for some, but worse for others than contractarianism.⁴⁶ Those sentient beings who are excluded from rights by social contract theory would gain the right to have their interests taken into equal account in utilitarian calculations. And those sentient beings who received minimal rights from contractarianism may be better protected by utilitarianism’s right to equal consideration in utilitarian calculations. But for those who receive the “full” package of rights from contractarianism—typically able humans—the utilitarian right to equal consideration is a net loss: while others might occasionally be called by utilitarianism to provide you with greater positive aid, you risk being called on to sacrifice your happiness for others. If, for instance, ten people could be saved if you were killed, utilitarianism would have you killed. This observation is, of course, anything but new, but it does show that for many, utilitarianism would provide weaker rights than contractarianism.

46. For the sense of “right” used in this thesis, see section 1.1.

Moreover, because utilitarianism allows the costs imposed on duty-holders to be balanced out by gains to others, utilitarianism is much more onerous from the perspective of duty-holders. Because contractarianism requires the costs imposed on duty-holders to be outweighed by benefits *to those duty-holders*, contractarianism must be less onerous than utilitarianism. Thus contractarianism takes better account of our dual roles as right-holders and duty-holders.

Thus, in assessing utilitarianism's performance as it compares to contractarianism, we find that the set of duty-holders to whom utilitarianism can justify duties will be comparably limited by its appeal to the three fundamental undefended premises. Moreover, because utilitarianism calls for duty-holders to make uncompensated sacrifices, *akrasia* will pose a much bigger problem for utilitarianism, weakening the strength of its argument.⁴⁷ The utilitarian set of right-holders is much broader than the contractarian set, making it perhaps more than sufficient for most people, but also leading it to include somewhat threatening right-holders who call for uncompensated sacrifices by humans. And while the utilitarian right to equal consideration may be more generous from the perspective of those excluded from moral consideration or offered minimal rights by contractarianism, from the perspective of those who would be full right-holders in a contractarian scheme, the utilitarian set of rights is less generous: one is entitled to equal consideration, but one does not have strict rights protecting one's important interests. Moreover, the utilitarian set of rights is much more onerous from the perspective of duty-bearers than the contractarian set.

47. When I claim that utilitarianism calls for uncompensated sacrifices, I mean that it may call for a duty-holder to sacrifice utility so that someone else might benefit, though the duty-holder will not benefit from that sacrifice.

5.2.2 Kantianism

Duty-Holders and Arguments

Regan contends that all beings who satisfy the “subject-of-a-life criterion” have inherent moral value, and that it is wrong, other things being equal, to harm or use as mere means any individual with inherent moral value.⁴⁸ Regan’s argument grounds these claims in our considered judgments and the principles that best explain them. Aside from a host of procedural and formal requirements for justifications of moral claims—like that we reason coolly and rationally, and that the moral principles we develop be consistent—the only consideration that Regan appeals to is fit with our considered intuitions in reflective equilibrium.⁴⁹ Thus, the neo-Kantian argument is in a worse position than the utilitarian argument with respect to the set of duty-holders. Singer made a serious attempt to get beyond moral intuitions, though I have questioned the success of his attempts. Regan relies much more heavily on much more demanding moral intuitions, for instance, that it is wrong to harm any subject of a life, or to use it as mere means.⁵⁰ Insofar as Regan relies exclusively on demanding moral intuitions, his argument can only address a very limited set of duty-holders.⁵¹

48. Regan, *The Case for Animal Rights*.

49. *ibid.*, 126–135, 146.

50. See section 4.2.

51. Regan considers the objection that by relying so heavily on considered intuitions and reflective equilibrium, he risks failing to generate *agreement* between various parties, instead leading to a different set of moral principles for each moral agent. But Regan’s response reveals that he completely failed to grasp the force of the objection: “What we can claim, with good reason, is that principles we validate in the manner outlined in the preceding are those that *would* be accepted by all who make ideal moral judgments about which principles to accept. What we can claim, in other words, is *a consensus among all ideal judges* concerning what principles are binding on all” (139). He attempts to explain disagreement by appeal to human fallibility, but he has failed to offer a procedure for determining whose starting intuitions are false and whose are correct, or for reconciling different starting intuitions, or for reconciling the different considered judgments reached by different people adjusting their starting intuitions and the principles explaining them randomly until they achieve a fit (139).

While the considered judgments may not be false, Regan's reliance on these premises and his inability to justify them limit his theory's ability to generate security through reasoned argumentation. Once again, it seems that the people against whom we most need security are not included in the set of duty-holders that Kantianism is capable of addressing.

As with utilitarianism, *akrasia* will be a bigger problem for Kantianism than for contractarianism insofar as Kantianism calls on duty-holders to make uncompensated sacrifices.

Right-Holders

Kantian theories include in the set of right-holders all and only those who meet the "subject-of-a-life criterion."⁵² The "subject-of-a-life criterion" entails that individuals have rights if and only if they have:

beliefs and desires; perception, memory, and a sense of the future, including their own future; an emotional life together with feelings of pleasure and pain; preference- and welfare-interests; the ability to initiate action in pursuit of their desires and goals; a psychophysical identity over time; and an individual welfare in the sense that their experiential life fares well or ill for them, logically independently of their utility for others and logically independently of their being the object of anyone else's interests.⁵³

52. Regan, *The Case for Animal Rights*, 243. This point could have been made in a more complicated way that best reflects what Regan says—*viz.* that all those who have inherent value are right-holders, and all those who meet the "subject-of-a-life criterion" have inherent value—but, as I argued in section 4.2, this complication is simply distracting and does not add anything relevant.

53. *ibid.*

Thus, Kantian theories generate a set of right-holders that is somewhat less comprehensive than that generated by utilitarianism—at least some sentient individuals would not meet all of the requirements of the “subject-of-a-life criterion.” One can note that some marginal human cases would not meet these requirements either: human infants and severely cognitively disabled humans would lack many of the cognitive requirements.

Is the Kantian set of right-holders sufficiently broad? Again, this is a subjective question. As with utilitarianism, it is likely that many Kantians would subjectively be satisfied with a narrower set of right-holders. And it seems unlikely that many people will find the Kantian set to be insufficiently broad, though there may be a few who do (especially people who feel strongly about sentient non-subjects-of-a-life, or about insentient nature).

How stable is membership in the Kantian set of right-holders? Membership, as defined by the subject-of-a-life criterion, is less stable than membership in utilitarianism’s sentience set: insofar as the former adds requirements to the latter, this must logically be the case. Relative only to contractarianism’s set of right-holders, as defined by the ability to impose significant costs on others in response to ill treatment, membership in the Kantian subject-of-a-life set is more stable: it is much easier to lose one’s ability to impose costs on others than to lose the abilities required to be subject-of-a-life. However, contractarianism does insure against this instability with protectorate status: insofar as we all recognize that our membership is unstable, we should all recognize that it is in our interest to insure against our loss of status by extending protections to humans beyond the set of right-holders. With this modification, it is possible that contractarianism offers more stable membership than

Kantianism. If all people who qualify for rights according to contractarianism desire that they should be protected even if they should no longer be subjects-of-lives, then this is sufficient for the establishment of an insurance-based protectorate status for all humans, even under conditions where they are no longer subjects of lives. Is it plausible that people have such a desire? That is certainly questionable, but this leads to a further point: it seems that contractarianism will generate precisely the kind of stability we want. Insofar as contractarianism's insurance-based protectorate status is based precisely on our desires to insure ourselves against certain conditions, and perhaps not to insure ourselves against others, contractarianism seems bound to produce precisely the kind of stability we want for ourselves.

There is one important limitation to contractarianism's ability to produce the kind of stability *we* want: such protectorate status arises from our unanimous agreement in interests. Should there be profound disagreement over whether or not humans who are not subjects-of-lives should be protected, those who want protection in such cases will not get it from contractarianism. Does this make contractarianism more or less stable than Kantianism? Contractarianism offers membership in the set of right- and protection-holders that is at least as stable, if not more so, than the Kantian view: because we are evaluating theories by their ability to secure our interests through reasoned argumentation, all theories need to secure a level of unanimity among duty-holders. Insofar as contractarianism can incorporate all protections on which there is unanimity, it is optimally placed to guarantee the stability of our membership, even should we no longer be subjects-of-lives, so long as we are unanimous in wanting such protection.

Does Kantianism include threatening right-holders? The subject-of-a-life set is similar to the utilitarian sentience set, but less comprehensive because of the added requirements. Thus, the Kantian set will be less threatening than the utilitarian set; however, it may still be threatening in a manner similar to the utilitarian set, if less significant in degree. While it is not seriously threatening to the ability for humans to survive, it will still call for duty-holders to make uncompensated sacrifices. Thus, it will require human duty-holders to sacrifice some of their interests. Because contractarianism rules out the possibility of uncompensated sacrifices on the part of duty-holders, the Kantian set will be more threatening than the contractarian set.

Rights

Regan's theory leads to negative rights for all beings who meet the subject of a life requirement not to be harmed, not to be discriminated against arbitrarily, and not to be used as a mere means.⁵⁴ For most, the Kantian set of rights is more generous from our perspective as right-holders than the contractarian set of rights. But from the perspective of duty-holders, while the Kantian set of duties is far less demanding than the utilitarian set, it is much more onerous than the contractarian set. Because Regan's approach focuses exclusively on the perspective of right-holders, ignoring the perspective of duty-holders, it leads to more onerous duties than contractarianism. Because contractarianism considers our role as both right-holders and duty-holders, it is in a better position to balance our competing interests than Kantianism.

So Kantianism offers a much weaker intuition-based argument that can appeal to a much narrower set of duty-holders than contract theory, and that faces a greater *akrasia* problem. The Kantian set of right-holders is broad, but no more stable than

54. Regan, *The Case for Animal Rights*, 187, 248. For more on these rights, see section 4.2.

the contractarian set when we consider protection-based insurance against disability. The Kantian set of right-holders is also more threatening than the contractarian set, though less threatening than the utilitarian set. Finally, while the Kantian set of rights is more generous, it is also much more onerous because Regan disregards the perspective of duty-holders.

5.2.3 Capabilities Theory

As explained in section 4.3, there are (at least) three possible interpretations of Nussbaum's capabilities approach, each of which affects the specification of the right-holders and the argument. The "*wonder* interpretation," I have contended, best fits with the arguments that Nussbaum develops, and claims that all individuals who inspire wonder are owed rights because they inspire wonder. The "*overlapping consensus*" interpretation, on the other hand, grounds the argument in the Rawlsian idea of an overlapping consensus, which may limit the set of right-holders. The "*sentience* interpretation" is neutral between the two previous interpretations of the argument, but limits the set of right-holders to sentient beings. While I have contended that the *wonder* interpretation is most consistent with Nussbaum's arguments, at various times she explicitly endorses the other two interpretations. In this evaluation of Nussbaum's capability approach, I will consider how each of these interpretations fares according to the element-specific desiderata.

Duty-Holders and Arguments

The primary argument that Nussbaum offers for her capabilities theory appeals to wonder, which is why I have emphasized the *wonder* interpretation over the *overlapping consensus* interpretation and the *sentience* interpretation. Thus, Nussbaum begins with Aristotle's claim "that there is something wonderful and wonder-inspiring in all the complex forms of life in nature."⁵⁵ Unlike Aristotle, Nussbaum derives an ethical judgment from the aesthetic judgment:

[I]f we feel wonder looking at a complex organism, that wonder at least suggests the idea that it is good for that being to persist and flourish as the kind of thing it is. The idea is at least closely related to an ethical judgment that it is wrong when the flourishing of a creature is blocked by the harmful agency of another. That more complex idea lies at the heart of the capabilities approach.⁵⁶

Nussbaum concludes:

[The capabilities approach] goes beyond the contractarian view in its starting point, a basic wonder at living beings, and a wish for their flourishing, and for a world in which creatures of many types flourish.⁵⁷

These passages reveal precisely how narrow an audience the *wonder* interpretation is capable of addressing. Firstly, the approach can only address those who start with the judgment that all complex forms of life are wonderful. Secondly, the approach can only address those who are willing to infer the relevant ethical judgment from the first

55. Nussbaum, *Frontiers of justice*, 347.

56. *ibid.*, 349.

57. *ibid.*

aesthetic judgment. Nussbaum appears to acknowledge that this inference does not follow deductively, using terms like “suggests” and “closely related.” While Nussbaum is right to suggest that the capabilities approach “goes *beyond* the contractarian view *in its starting point*,”⁵⁸ this feature is a weakness of the capabilities approach when evaluated from a security perspective. The starting point—the assumptions that the audience must endorse to find the capabilities theory compelling—is already quite demanding and significantly narrows the set of duty-holders to whom capabilities theory is capable of justifying obligations. Of all the theories considered in this thesis, the capabilities approach is capable of addressing only the narrowest audience.

By grounding the duties and rights of capabilities theory in the aesthetic judgment that something is wonderful, capabilities theory also risks giving duty-holders comparably weak reasons for fulfilling their duties. It is beyond the scope of this thesis to adequately address the relation between aesthetic and moral judgments. Instead, I seek only to suggest that we often sacrifice our aesthetic values for the sake of other values. I believe that most of us rank our own well-being, and that of those we care for, above our aesthetic values. We would be willing to sacrifice the latter to pursue the former. For instance, while we may find many low-cost forms of high-density housing to be aesthetically unappealing, we would consider it inappropriate to prohibit such developments. And if we faced a choice between burning a priceless work of art or freezing to death, the vast majority would choose to burn the art. Our

58. Nussbaum, *Frontiers of justice*, 349. My emphasis.

aesthetic interests are rather easily outweighed by competing interests.⁵⁹

Right-Holders

It is unclear who would be included in the set of right-holders according to capabilities theory. Nussbaum explicitly claims that all and only sentient beings are included as right-holders (the “*sentience* interpretation”), but this specification appears arbitrary as it has little connection with the grounds for capabilities theory.⁶⁰ Two alternative answers stand out as plausible: all individuals who inspire wonder, or all those on whose membership an overlapping consensus can be reached.⁶¹ As I argued in the section 4.3, Nussbaum’s appeal to the notion of an overlapping consensus does not fit well with the arguments she presents. The answer that fits best with Nussbaum’s arguments is that individuals that inspire wonder are right-holders. If we accept the *sentience* interpretation, then the capabilities theory will generate the same set of right-holders as utilitarianism, including the vast majority of animals and marginal cases. If, instead, we go with the *overlapping consensus* interpretation, the set of right-holders will be specified by an account that can be articulated independently of any substantive metaphysical or epistemological doctrines or conceptions of the good.⁶² It is unclear who would be included on such a view; we could maybe accept the *sentience* criterion as an object of overlapping consensus—it would generate what

59. Although I must confess that there are certainly counterexamples to my claim. At the time of writing, the capture of the UNESCO World Heritage Site at Palmyra has received considerably more media attention than the capture of the modern city of Palmyra, prompting Dr. David Roberts of King’s College, London, to contribute a section entitled “People v Statues” to a BBC article on the topic. Roberts responds to the worry that the media is paying more attention to the loss of artefacts than to the human cost of the capture of Palmyra by noting that the concerns are not mutually exclusive. “Palmyra: Islamic State locks down ancient city’s museum,” BBC, 2015, <http://www.bbc.com/news/world-middle-east-32859347>.

60. Nussbaum, *Frontiers of justice*, 361–362.

61. See section 4.3. See also *ibid.*, 93–94, 163, 346–347.

62. *ibid.*, 163.

is likely to be the most extensive set of right-holders that could possibly be the object of an overlapping consensus, so it is the most charitable assumption we could make. Finally, if we go with the *wonder* interpretation, as I suggest, then the set of right-holders will include all individuals who are sources of wonder (setting aside the Euthyphro problem of whether such individuals must *have* some property that makes them wonderful, or must *be found* wonderful by some relevant set of people).⁶³ It is unclear who would be a right-holder on this interpretation. Are *all* humans sources of wonder? That seems unlikely. Mountain ranges, storms, and ecosystems, to consider a few examples, seem like appropriate sources of wonder, so they would likely be included.⁶⁴

Is the capabilities set of right-holders sufficiently broad? If we take the *sentience* interpretation, then the answer to this question will be the same as the answer for utilitarianism, namely, that the set is broader than is likely necessary for the vast majority of people, but perhaps too narrow for a few who love insentient things. And I have suggested accepting the sentient set as the most charitable result of the *overlapping consensus* interpretation, so the answer there remains the same. Finally, if we accept the *wonder* interpretation, which I suggest fits best with Nussbaum's arguments, then it is highly unlikely that we could find anyone who would find the

63. "Is the pious being loved by the gods because it is pious, or is it pious because it is being loved by the gods?" Plato, "Euthyphro," in *Plato: Complete Works*, ed. John M. Cooper, trans. G.M.A. Grube (Indianapolis: Hackett Publishing Company, 1997), 9 (10a).

64. Nussbaum often limits the range of sources of wonder she designates to "complex natural organisms" (94), "living creatures" (328), "the complex forms of life in nature" (347), "complex organism" (349), or "living beings" (349). Nussbaum, *Frontiers of justice*. The inconsistency makes it unclear whether these qualifications are added necessary conditions, so that a being is worthy of moral protection only if it is both wonderful and alive/complex/natural, or whether they are sufficient conditions for being wonderful, and thus worthy of moral protection. If it is the former, then my examples will not do: they may be wonderful, and some of them are complex and natural, but (arguably) they are not alive. Cancer, weeds, pests, and violent animal predators might meet all of the qualifications, be wonderful in certain senses, and yet strike us as inappropriate candidates for moral protection.

set of right-holders insufficient, though the vast majority would be satisfied with narrower sets.

How stable is membership in the set of right-holders according to capabilities theory? Once again, the *sentience* interpretation is equivalent to utilitarianism in this respect: membership is quite stable. Although I have proposed that the *overlapping consensus* interpretation might also yield the sentience set, the stability of membership on this interpretation may not be equal to that of the *sentience* interpretation insofar as the *overlapping consensus* interpretation relies on the stability of the consensus. Should the consensus change over time, as members change, and as world views and conceptions of the good change, then membership in the set of right-holders on this view may prove less stable than on the sentience interpretation. Alternatively, membership could be more stable—it could be more likely that one should lose one’s sentience than that the consensus should shift to exclude one over the course of one’s lifetime.⁶⁵ I do not know how to evaluate which is more likely, but I hope that the difference is sufficiently insignificant that it can be set aside. The stability of one’s membership in the set of right-holders is less clear still according to the *wonder* interpretation. The Euthyphro problem becomes important here: if an object’s being “wonderful” relies exclusively on people’s evaluations of the object, then it seems that membership is unstable—especially if the wonder people take in the object diminishes over prolonged exposure to it. If, instead, being “wonderful” relies on having some property, then membership will be as stable as the object’s having that property—it could be easy to lose (consider the blossoming flowers of a magnolia tree) or very difficult to lose (consider the complicated root structure of the

65. If we take “overlapping consensus” to be an abstract and universal theoretical concept that does not depend on the attitudes or views of particular individuals in particular times, this problem does not arise. However, it is not obvious that this universal interpretation is a live possibility.

magnolia tree, or its ability to photosynthesize). Thus, determining the stability of membership according to the *wonder* interpretation is quite difficult without a careful specification of the notion of “wonder.”

Whether capabilities theory extends rights to threatening right-holders depends heavily on the interpretation endorsed. The *overlapping consensus* interpretation is least likely to include threatening right-holders insofar as the consensus appeals to the interests of the parties forming the consensus. The notion of an overlapping consensus would seem to exclude the possibility of reaching an overlapping consensus on a policy that would threaten humankind. The *sentience* interpretation is equivalent to utilitarianism with regard to threat: it is not severely threatening, but does call for uncompensated sacrifices, thereby threatening the interest of duty-holders to some extent. The *wonder* interpretation is the most likely to contain threatening members. Indeed, some things inspire wonder because of their threatening nature: polar and grizzly bears are *aweful* (in both the current and the antiquated senses of the term) because of their fierce natures, their strength, and what they could do to us.⁶⁶ Fire, seas, and violent storms may inspire wonder for the same reason. Insofar as all of our sources of food, shelter, and clothing inspire wonder, the source of wonder set is immensely threatening to humans. Likewise, insofar as things that are dangerous to humans inspire wonder, the source of wonder set is threatening to humans. These claims depend, of course, on the rights extended to these threatening members: should the rights extended to them be limited by the condition that human security is paramount, then the theory is less threatening to humans. But some wonderful

66. One definition of “awful” from the Oxford English Dictionary is particularly telling in combining the kind of wonder that calls for respect with fear: “Worthy of, or commanding, profound respect or reverential fear.” (“awful, adj.,” in *OED Online* (Oxford University Press, 2015), <http://www.oed.com.proxy.queensu.ca/view/Entry/13943?redirectedFrom=awful&>).

grounds would need to be offered for preferring human security in this way.

Rights

Capabilities theory generates a very generous set of rights, but disregards the perspective of duty-holders. The set of rights generated is more generous than Regan's insofar as capabilities theory includes positive rights while Regan's includes only negative rights. And unlike utilitarianism, the rights generated by capabilities theory protect one from being made a sacrifice to the wellbeing of others. Finally, compared to the contractarian set of rights, the capabilities theory is *much* more generous.

However, the duties that capabilities theory gives rise to will be immensely exacting. The duties generated by the capabilities theory are clearly much more demanding than the duties of non-harm and respect that Regan calls for, and it is unclear how they compare to utilitarianism: while capabilities theory does not call on agents to sacrifice themselves without limit for the aggregate happiness of others, it may call for sacrifices that do not maximize net utility. For instance, while utilitarianism restricts harm to painful (or less pleasant) experiences, capabilities theory recognizes and proscribes non-painful harms, and would thus be more demanding in that regard. And capabilities theory is clearly more demanding than contractarianism from the perspective of duty-bearers insofar as contractarianism requires all duties to be to the advantage of the duty-bearers. So while the rights generated are certainly more generous than those generated by contractarianism, the duties are also much more onerous. What we want in terms of rights is an appropriate balance between our perspective as right-holders and our perspective as duty-holders, and capabilities theory fails to strike a balance between these two perspectives, heavily privileging the former

over the latter.

Thus, evaluating capabilities theory by these criteria is complicated by the three possible interpretations of capabilities theory. Its argument is strongly grounded in aesthetic intuitions, weakening the argument and limiting its ability to appeal to only a narrow set of duty-holders. The set of right-holders it generates is much broader than the contractarian set, and perhaps more stable, but risks including threatening members depending on the interpretation. Finally, while the rights generated by capabilities theory are much more generous than those of contract theory, they are also much more onerous from the perspective of duty-holders.

5.2.4 Contractarianism

Duty-Holders and Arguments

The set of duty-holders to which contractarianism can appeal is limited by the theory's account of human nature. Contemporary contractarians like David Gauthier appeal to a very thin account of human nature limited to an account of practical rationality as individual expected utility maximization. Hobbes and Hume, and generally those contractarians who attempt to derive theories of rights, appeal to a slightly thicker account of human nature, adding to the utility maximization account of practical rationality certain assumptions about human motivations.⁶⁷ Thus, Hobbes offered his physiological account of the source of desire, and used that to contend that all healthy humans desire those things that preserve their lives and loathe those things that harm their lives (I.vi., esp. I.vi.9-10). These contractarian theories are united in their endeavour to work from a very general account of human nature so as to

67. For further discussion of the account of human nature, see section 3.1.1.

maximize the scope of the set of duty-holders who will find the arguments persuasive. The goal is to rely on the thinnest characterization of the parties to the convention; the more (or stronger) assumptions one adds to the characterization, the more one risks diminishing the set of duty-providers who will find the arguments compelling. This does not mean that the thicker accounts are wrong, or that they are necessarily limited to a narrower audience, but only that the risk of excluding individuals from the set of duty-holders is higher the more assumptions we make about those duty-holders. As long as a thicker account accurately describes human nature, the problem of exclusion is avoided.

Nevertheless, exclusions there will be. Hobbes was aware of this, acknowledging that unhealthy individuals may have perverted natures that preclude them from being addressed by his arguments, though he may not have appreciated the full significance of the problem.⁶⁸ As long as there are exceptions to any helpful account of human nature, there will be exclusions from the contractarian theory. Moreover, as section 3.1.2 explained, some individuals will be excluded from the contract on the grounds that cooperation with them is a net loss to others. For instance, those with misanthropic desires will be excluded.

Not only does contractarianism seek to appeal to the widest possible set of duty-holders, but it attempts to give the strongest possible reasons for duty-holders to fulfill their duties. Contractarians endeavour to show that obeying the moral code will make one's life go best according to one's own values.⁶⁹ The corollary is that failing to fulfill one's duty will make one's life go worse. Beyond the formal account

68. See Hampton, "Hobbes and Ethical Naturalism." Hampton argues that Hobbes's appeal to health threatens his naturalistic commitments. The contractarian account developed here excludes individuals from the contract not on the grounds that they are unhealthy, but on the grounds that cooperation with them on terms that everyone could accept is impossible. See section 3.1.2.

69. See section 3.1.4.

of practical rationality, the contractarian theory used here appeals to the desire for one's continued life and health. Taken together, the desires for continued life and health, and for as many desires to be satisfied as possible, are not easily outweighed. We rarely sacrifice these interests for the sake of aesthetic values, for instance, and many of us sometimes allow these interests to lead us to do what we ourselves may consider to be morally bad, including harming others. Insofar as contractarianism successfully grounds moral principles in these interests, it offers the strongest reasons for duty-holders to fulfill their duties.

Right-Holders

According to contractarianism, the set of right-holders will include those who can respond to the unrestrained behaviour of others by imposing costs that outweigh the benefits of such unrestrained behaviour.⁷⁰ As I have argued, this means that some animals will be included in the set of right-holders, depending on their abilities and on the benefits that humans derive from them.⁷¹ Contractarianism also generates protections for human infants, disabled and infirm humans and various animals.⁷² The contractarian set of right-holders is less comprehensive than (and contained by) the utilitarian set, the *sentience* interpretation of the capabilities theory, and likely also the Kantian set. The relations between the Kantian set, the *overlapping consensus* and *wonder* interpretations of the capability set, and the contractarian set are more complicated than the straight-forward containment relation, but it is most charitable to assume that the contractarian set is less comprehensive than those alternatives.

Is the contractarian set of right-holders sufficiently broad? I have argued above

70. See section 3.1.5.

71. See section 3.2.

72. See sections 3.1.7 and 3.3.

that other theorists will tend to be satisfied with less comprehensive sets of right-holders than their preferred theories call for, and this may lead many people to be satisfied with the contractarian set.⁷³ Insofar as our partial sympathies are mostly focused on the humans with whom we keep company and to our companion animals, the contractarian set of right-holders will be judged by many to be sufficient. Nevertheless, there may be contractarians (among others) who find the contractarian set of right-holders to be insufficiently broad. For instance, a contractarian who shared the judgments of wonder to which capabilities theory appeals may find the exclusion of many natural entities from the set of right-holders to be cause for sadness and regret. Because the question is subjective, a definitive argument is not available here, and empirical evidence would be needed to prove trends, but I suspect that Mill and Hume were right that most people's sympathies are mostly limited to their friends, family, acquaintances, and companion animals.⁷⁴ Insofar as this is true, most people will judge the contractarian set of right- and protection-holders, which includes all of the immediately aforementioned, to be sufficient.

How stable is membership in the contractarian set of right-holders? When we focus exclusively on the set of right-holders, membership in the contractarian set is the easiest to lose of all the competing theories considered here. Contractarianism insures against this instability by appealing to protectorate status: insofar as we all recognize that our membership is unstable, we should all recognize that it is in our interest to insure against our loss of status by extending protections to humans beyond the set of right-holders. With this modification, I think that it becomes less clear whether membership in the set of right-holders (including protectorates) is more

73. See section 5.2.1.

74. Hume, *An Inquiry Concerning the Principles of Morals*, 19, 55; Hume, *A Treatise of Human Nature*, (III.ii.1) 532–536; Mill, “Utilitarianism,” 195.

or less stable for contractarianism than for the competing theories. Membership in the contractarian set will be less stable for non-humans: their membership is heavily dependent on what benefits they can offer humans, their abilities to withhold these benefits or to inflict harm, and the benefits that humans can derive from mistreating animals. But human membership may be at least as stable, if not more so, than membership according to other theories. Thus, while contractarianism may not be the best theory when it comes to stability of membership in the set of right-holders, it is unclear whether or not contractarianism is the worst. For humans, at least, and thus for duty-holders, membership is quite stable.

However, there is another way to lose one's membership in the contractarian set of right-holders that does not involve any change in one's own abilities. The contractarian theory only applies to the circumstances of justice which allow for positive-sum interaction. Should the circumstances of justice fail to obtain, not only would one lose one's membership in the set of right-holders, but the whole contractarian moral scheme would collapse. There is little to be said in response to this matter: it is hoped that the circumstances of justice are unlikely to cease to hold, and there is also considerable reason for many to take great pains to ensure that they are preserved insofar as their preservation is within our control. Moreover, given the considerable shifts in public opinion on matters like slavery, segregation, homosexuality, and even the moral status of animals, it may be more likely that the moral intuitions on which other theories are grounded will shift than that the circumstances of justice will cease to hold in a dramatic way. But the prospect of such shifts should trouble us—particularly, the specter of such drastic climate change that moderate scarcity should turn to radical scarcity.

Does contractarianism include any threatening right-holders? The very structure of the contractarian reasoning precludes this. Insofar as contractarianism insists that duties must be shown to be in the interest of the duty-bearers, it cannot show that duties are owed contrary to the interests of the duty-holders.

Rights

Other theories claim to generate a more generous set of rights than that generated by contractarianism. I will make the charitable (and plausible) concession that other theories likely generate more generous sets of rights than contractarianism. Once again, contractarians can (and likely ought) to recognize that it would be morally good if individuals were treated in the ways that other theories call for; what contractarians ought to object to is thinking that such behaviour is morally required. The contractarian can think that it is regrettable that a “nicer” set of rights cannot be justified by the contract method, yet maintain that trading the strength of the contract account for the nicer sets of rights is not a worthwhile trade.

While, from the perspective of right-holders, other theories yield more generous sets of rights, from the perspective duty-bearers, other theories are much more onerous. While we may like to have more rights, we may also regret being subject to more demanding duties. Contractarianism is more sensitive to the perspective of duty-bearers than any of the competing theories. One could argue that utilitarianism, insofar as it focuses on aggregate utility, will factor in the disutility to duty-bearers, but utilitarianism allows for the disutility to a duty-bearer to be offset by someone else’s utility. A duty-bearer will only find consolation in this offset if that duty-bearer cares as much about the utility of others as about his or her own. While fundamental

equality is a moral principle for utilitarianism, it is not a subjective principle of action, nor is it likely to accurately reflect the attitudes of many humans. When push comes to shove, most of us prefer our own welfare and that of those we care for over that of strangers and enemies. Contractarianism, on the other hand, demands that duties be in the interest of the duty-bearers themselves. Thus, contractarianism is more sensitive to the perspective of duty-bearers than utilitarianism, which is itself the most sensitive of all the competing theories to the perspective of duty-bearers.

Thus, contractarianism offers the strongest argument to the broadest set of duty-holders of all the theories considered. While the set of right-holders is not as broad as the other theories, many will deem it sufficient, and it is the least likely to contain threatening right-holders. Moreover, when protectorate status is factored in, membership in the set of right- and protection-holders is stable. Finally, while the rights generated are not as generous, the contractarian set of rights best balances our competing perspectives as right-holders and duty-bearers.

5.2.5 Final Evaluation

What we are left with is a question of balancing the weaknesses of the contractarian account of rights against the weaknesses of other accounts of rights. Competing accounts of rights may offer more generous sets of rights to a broader set of right-holders whose membership in that set is more stable, but many of us may be satisfied with more limited sets of right-holders. Moreover, the set of rights generated by competing theories may be *too* generous when assessed from the perspective of duty-holders. While competing theories may do better in these regards, contractarianism still does rather well: those we care about most tend to qualify for rights and protections which

balance our interests as right-holders and as duty-holders. But when it comes to giving the strongest reasons to the broadest set of duty-holders, the competing theories do much worse than contractarianism. Other theories do not pay sufficient attention to duty-holders: they focus exclusively on the interests that individuals have as right-holders, ignore the scope of the audience they can reach, and offer comparably weak reasons for duty-holders to comply. Singer is the exception in this regard, though I contend that because of the relative weakness of his arguments and the call for duty-holders to make uncompensated sacrifices, his theory will be much less broadly persuasive than contractarianism.⁷⁵

I think that, from a security perspective, contractarianism stands out as the best theory of rights, but I must acknowledge that this argument is not deductively valid. For the argument to be valid, some sort of dominance argument would need to apply, or else a complete ranking of the various desiderata would be needed. While I originally hoped that a dominance argument would be available, it seems clear that, although contractarianism does well with respect to all of the desiderata, on some of the desiderata other theories do better.⁷⁶ At the same time, a complete ranking of the desiderata seems impossible to articulate. At most, we can say that all of the desiderata are important, so a theory that failed in one respect would be worse than a theory that did acceptably in all of the desiderata. For instance, a theory that offered the most sufficient set of right-holders but that failed to address any duty-holders

75. For my assessment of Singer's argument, see section 5.2.1.

76. There are two kinds of dominance argument that I am considering. Theory A *strongly* dominates theory B when, by all criteria, A is better than B. Theory A *weakly* dominates theory B when (a) by all criteria, A is at least as good as B, and (b) by at least one criterion, A is better than B. As soon as B is better than A in at least one criterion, it is no longer possible for A to dominate B, weakly or strongly. The concepts of weak and strong dominances are common to the study of rational choice theory. See, for instance, Martin Peterson, *An Introduction to Decision Theory* (Cambridge, UK: Cambridge University Press, 2009), 41–43.

would be a complete failure. Further precision in the ranking appears impossible.

Do the competing theories fail completely at giving strong reasons to duty-holders? That seems unfair and impossible to assert with confidence: how many people would have to be excluded for a theory to count as failing in this task? A host of antecedent value judgments would need to be made answering questions of the form “Would sacrificing x for the sake of increasing the set of duty-holders by y members be worthwhile?” And even with non-question-begging answers to these questions, we would then need to know how many people indeed endorse which assumptions. I think other theories do quite poorly with respect to giving the strongest reasons to the broadest set of duty-holders, with the exception of Singer, but I am not confident that they fail, nor that a non-question-begging argument can be produced to show that the sacrifices they make on this front are not compensated for by the gains they make on other fronts.

But this weakness of the competing theories is still important: the distinctive aspect of basic rights—as opposed to other moral tools and concepts like goodness, badness, or flourishing, and as opposed to other kinds of rights like rights acquired by contract or by position within an institution—is that such rights call for everyone’s respect. Speaking to the broadest possible set of duty-holders is key to a theory of rights as opposed to a theory of the good life, or a complete account of moral goodness and badness. While the argument above may not show that, so long as all of the premises are true, the conclusion that contractarianism is the best theory from a security perspective follows *necessarily*, the argument still favours that conclusion.

Not only does the greater exclusivity of the alternative theories make them less fitting *qua* theories of rights, it also makes them weaker from a security perspective. As

discussed in section 3.1.2, even contractarianism will exclude some individuals from the contract, but the costs of that exclusion—the costs of detection and punishment, prevention, precedent, and foregone benefits of compromise—are always significant. One way to refocus the disagreement between contract theory and the alternative theories considered here is to consider it as a disagreement over the value of coalitions of various sizes. The contract theorist judges the broad coalition it forms to be of higher value than both the broader coalitions with weaker rights and the narrower coalitions with stronger rights. Proponents of the alternative theories here, on the other hand, judge their narrower coalitions to be worth more than the broader contractarian coalition with weaker rights.

The solution to this valuation problem that I contend would be optimal is a form of mixed strategy. Talk of rights and duties should be reserved for the theory with the broadest and strongest appeal—contractarianism—while the other theories should be presented as theories of moral goodness and badness—theories of what we ought to do, but what we can neither be coerced to do nor punished for not doing. Because contractarianism is capable of addressing the broadest audience, it is best placed to function as a theory of rights. And the contributions to security that other theories can make may not be diminished by representing them in terms of other moral concepts like moral goodness and flourishing: insofar as people endorse the assumptions of those theories, they should follow their theories so long as doing so does not violate anyone's contractarian rights. This solution allows us to combine the best elements of all of the theories: our rights are secure against the broadest possible set of duty-holders, and other theories can make gains above this threshold.

One implication of this approach which assigns rights and duties to the theory that

best ensures our security, while leaving other theories to discuss moral goodness and badness, is that we may place less emphasis on accounts of rights. Theories of rights do not provide us with complete accounts of the good life, or with all of the guidance we could possibly need to lead good, moral lives. They are a tool for our security, and that is all. Perhaps we have been stretching our theories of rights too far, trying to provide a complete account of morality in terms of rights. On my proposal, theories of rights would be very important for our security, but they would not answer all of the questions we might have about how to behave. To answer those further questions, we could appeal to other theories more freely, without feeling the need to bring everyone with us. We could say, "This theory is good for me," without feeling the need to prove that it is good for everyone. With this division effected, it will be found that contractarianism is not incompatible with any of the other theories considered here. Contractarianism's strength is its ability to offer the most persuasive account of minimal obligations of justice. This account is built on minimalist assumptions, but the other theories provide a more complete account of what it is to act morally and to live well to people who endorse further moral assumptions.

Chapter 6

Conclusion

In this thesis, I have argued that mutual advantage contract theory offers us the best theory of moral rights from a security perspective, and that this theory entails rights and protections for various animals and marginal cases. This is, admittedly, a narrow perspective: security is just one thing that we might want from a moral theory. Nevertheless, I think that security is a very important desideratum of a moral theory, and is particularly important for theories of moral *rights*, as opposed to other moral concepts.

According to contract theory, rights would be held by those animals and marginal cases who are (a) capable of responding appropriately to the behaviour of others (b) by imposing costs in response to unrestrained behaviour (c) which outweigh the benefits that others derive from their unrestrained behaviour. Being capable of imposing cost alone is not sufficient to qualify for rights—one must be able and willing to impose costs in response to undesired treatment, and to refrain from imposing costs in response to desired treatment. The sense of “costs” here is that of opportunity cost: benefits that would be conferred in response to desired treatment but withheld in response to undesired treatment are “costs” in the relevant sense just as much as

harms that would be conferred in response to undesired treatment but withheld in response to desired treatment.¹

Applying these criteria to animals to determine which animals qualify for which rights remains a complicated affair. I offered two distinct ways of categorizing animals for the sake of generating further guidance concerning their moral rights: by the kinds of benefits they bestow on us, and by species.² Based on the benefits bestowed, I argued that animals used for companionship, security, hunting assistance, transportation, entertainment, medical service, nourishment, and clothing or other accessories are owed basic rights to adequate food, health care, and protection from predators and the elements. Animals used for companionship are owed further rights, which must be negotiated on a case by case basis between the animals and their owners. Animals used for security, hunting assistance, transportation, entertainment, and medical service may be owed further rights if they withhold the benefit in question unless certain additional conditions are met, and if those conditions are not more costly than the benefit derived. Animals used for medical testing or hunting have no rights: the benefit derived requires that the animals' interests be disregarded.³

Based on the categorization according to species, I argued that dogs should have rights against mental and physical abuse and against use in 'frivolous' medical experiments, and rights to adequate nourishment, health care, and protection from predators and the elements, unless they are serving in medical experiments. Dogs may also receive further rights, as negotiated between dogs and their owners. Cows

1. For the full argument supporting these conclusions, see section 3.1.5.

2. Or, rather, insofar as the species I selected were meant to be representative, we could say categories of species. But caution is needed: the rights generated by species were already suggestive at best due to wide variations in abilities among species-members, and we can anticipate that extending the rights to similar species will be even more fraught by variations in abilities.

3. For the full argument, see section 3.2.1.

should have rights to adequate food, health care, and protection, but bears and squirrels do not receive any rights.⁴ From these species, we may, with caution, generalize from dogs to domesticated animals, from cows to farm animals, from bears to wild animals, and from squirrels to liminal animals.⁵

Contract theory also generates a form of indirect moral consideration that I called protectorate status.⁶ I argued that mutual advantage contract theory would support rather strong and universal protections for those wild animals required for the continuation of human life, including those animals and ecosystems whose destruction would contribute to climate change. These protections would include protections against negative human interference, and likely an entitlement to beneficial human intervention. Likewise, protections for all animals against cruelty, or harm disproportional to whatever benefit is to be gained thereby, are likely to arise from human disgust at the thought of cruelty to animals. Finally, protections for pets against harm at the hands of anyone other than their owners, and for certain wild species could be generated, though because the case for these protections appeals to somewhat weaker and less universal preferences, the protections may be weaker.

I also argued that mutual advantage contractarianism would afford protections to disabled and infirm humans based on a kind of insurance reasoning: insofar as we are all highly likely to experience periods of disability over the course of our lifetimes, we all stand to benefit from protections for disabled and infirm humans. And I argued that parents would insist on protections for their children against others, and that communities would insist on protections for children against their parents.⁷ Thus,

4. However, as noted in footnote 68 on page 125, the right to nourishment becomes problematic when the interests of the cow (or other animal) diverge from the interests of its owner.

5. For the full argument supporting these conclusions, see section 3.2.2.

6. For the full definition of protectorate status, see section 3.1.7.

7. See section 3.1.7.

when considering rights and protections together, mutual advantage contract theory is more than capable of extending moral status to animals and marginal cases.

After developing the contract theory of rights for animals and marginal cases, I argued in chapter 5 that contract theory offers the strongest account of rights from a security perspective compared to utilitarianism, Kantianism, and capability theory. From a security perspective, theories of rights are judged by their ability to convince duty-holders not to interfere in our pursuit of what we desire, with particular emphasis on our desire for continued life and health.⁸ I applied this overarching aim to each of the four elements of a theory of rights that I identified: the set of duty-holders against whom rights apply, the argument convincing duty-holders to respect rights, the set of right-holders protected by rights, and the set of rights protecting them. In this manner, I derived the following security-based criteria-specific desiderata by which to assess rights:

1. *Ceteris paribus*, the theory with the broadest set of duty-holders is best.
2. *Ceteris paribus*, the theory with the argument that addresses the broadest set of duty-holders is best.
3. *Ceteris paribus*, the theory with the argument that gives duty-holders the strongest reasons—*viz.*, the least easily outweighed or overridden reasons—for respecting rights is best.
4. *Ceteris paribus*, the theory that includes the most individuals whose security we desire (including ourselves) in the set of right-holders is best.

8. For the full definition and defence of security as the central desideratum, see section 5.1.1.

-
5. *Ceteris paribus*, the theory which makes membership in the set of right-holders the most stable—*viz.*, the least easily lost—is best.
 6. *Ceteris paribus*, the theory that includes the fewest threatening right-holders is best.
 7. *Ceteris paribus*, the theory that best balances our interest as right-holders in having as generous a set of rights as possible with our interest as duty-holders in avoiding overly onerous duties is best.⁹

I then evaluated each of the competing theories, in turn, according to these criteria. Based on that analysis, I argued that, while the competing theories may offer more generous sets of rights to broader sets of right-holders whose membership is more stable, contractarianism nevertheless does well (even if not best) in these regards. Moreover, the set of right-holders generated by the other theories may be more generous than many of us want, and the set of rights they generate may be far too onerous from our perspective as duty-holders, thereby failing to achieve an appropriate balance. Most importantly, I argued that contractarianism does immensely better than the other theories with respect to giving the strongest argument to the broadest possible set of duty-holders.¹⁰

For a deductively valid argument to show that the security-based desiderata entail that one theory is better than the others, either some form of dominance reasoning must apply, or else a complete ranking of the desiderata in terms of their relative

9. For the argument generating these desiderata, see section 5.1.3. While I distinguish the first and second criteria in this list, in my analysis of the competing theories I considered these criteria together because they amount to the same requirement.

10. For the complete argument, see section 5.2.

importance must be available.¹¹ Because of the mixed results above, dominance reasoning is unavailable. Moreover, I do not see how a complete ranking of relative importance is possible. The most that can be said is that a theory that failed completely in one of the criteria, but did well in others, would have failed completely. A theory that could generate wonderful rights but not for any right-holders, or could generate wonderful rights for all possible right-holders but could not justify them to any duty-holders, would fail from a security perspective. But it would be implausible to claim that the alternative theories considered here completely failed in any regard.¹²

Nonetheless, I contended that the argument, though not deductively valid, is strongly suggestive that contractarianism offers the strongest theory from a security perspective because it does well in all criteria, and because it drastically outperforms the other theories in addressing the strongest argument to the widest set of duty-holders. While the other theories do not fail to address arguments to duty-holders, they are very weak in this regard. This weakness entails that all of the costs associated with excluding individuals from cooperation identified in section 3.1.2 will be highest for the other theories—costs of detection and punishment, prevention, precedent, and foregone benefits of compromise. I contend that these costs are considerable from a security perspective, and suggest that contract theory is the strongest theory of rights.¹³

However, while I argue that contract theory offers the strongest theory of *rights*, I suggest that the best strategy is a mixed strategy: moral rights should be grounded in

11. See footnote 76 on page 212. There are two forms of dominance reasoning: A strongly dominates B when A is better than B in all criteria. A weakly dominates B when (a) A is at least as good as B in all criteria, and (b) A is better than B in at least one criterion.

12. For the complete argument, see section 5.2.5.

13. For the full argument supporting these conclusions, see section 5.2.5.

contract theory, while the other theories may provide the strongest accounts of other moral concepts like good, bad, flourishing, and virtue. Contract theory should not claim to tell us everything we could possibly want to know about morality in general; rather, it offers the strongest account of moral rights—basic moral entitlements—from a security perspective.¹⁴

Thus, I conclude that to answer the question “What moral obligations, if any, do we have toward animals?” we should consult the strongest theory of moral rights—the mutual advantage contract theory—and conclude that we must respect the contractual rights and protections of animals defended in this thesis.

14. For the complete argument, see section 5.2.5.

Bibliography

- “awful, adj.” In *OED Online*. Oxford University Press, 2015. <http://www.oed.com.proxy.queensu.ca/view/Entry/13943?redirectedFrom=awful&>.
- Barry, Brian. *A Treatise on Social Justice, Volume 1: Theories of Justice*. Los Angeles: University of California Press, 1989.
- Bernstein, Mark. “Contractualism and Animals.” *Philosophical Studies* 86 (1997): 49–72.
- Buchanan, James M. *The Limits of Liberty: Between Anarchy and Leviathan*. Vol. 7. Indianapolis: Liberty Fund, 2000.
- Cartwright, Nancy. *Nature’s Capacities and their Measurement*. Oxford: Clarendon Press, Oxford University Press, 1994.
- Cavalieri, Paola, and Will Kymlicka. “Expanding the Social Contract.” *Etica & Animali* 8 (1996): 5–33.
- Coeckelbergh, Mark. “Distributive Justice and Co-Operation in a World of Humans and Non-Humans: A Contractarian Argument for Drawing Non-Humans Into the Sphere of Justice.” *Res Publica* 15 (2009): 67–84.

- Cohen, Andrew I. "Contractarianism and Interspecies Welfare Conflicts." *Social Philosophy and Policy* 26, no. 1 (2008): 227–257.
- . "Contractarianism, Other-regarding Attitudes, and the Moral Standing of Nonhuman Animals." *Journal of Applied Philosophy* 24, no. 2 (2007): 188–201.
- DeGrazia, David. *Animal Rights: A Very Short Introduction*. Oxford: Oxford University Press, 2002.
- Dimock, Susan. "Contractarian Criminal Law Theory and *Mala Prohibita* Offences." Chap. 6 in *Criminalization: The Political Morality of the Criminal Law*, edited by R.A. Duff, Lindsay Farmer, S.E. Marshall, Massimo Renzo, and Victor Tadros. Oxford: Oxford University Press, 2014.
- . "Two Virtues of Contractarianism." *Journal of Value Inquiry* 37 (2003): 395–414.
- Donaldson, Sue, and Will Kymlicka. *Zoopolis: A Political Theory of Animal Rights*. Oxford: Oxford University Press, 2011.
- Driver, Julia. "A Humean Account of the Status and Character of Animals." In *The Oxford Handbook of Animal Ethics*, edited by Tom L. Beauchamp and R.G. Frey, 145–172. Oxford: Oxford University Press, 2011.
- Gaus, Gerald. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press, 2012.
- Gauthier, David P. *Morals by Agreement*. Oxford: Clarendon Press, Oxford University Press, 1986.

- Gauthier, David P. *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*. Oxford: Clarendon Press, Oxford University Press, 1969.
- Gert, Bernard. *Man and Citizen*. Atlantic Highlands, N.J.: Humanities Press, 1978.
- Hampton, Jean. "Hobbes and Ethical Naturalism." *Philosophical Perspectives* 6, *Ethics* (1992): 333–353.
- . *Hobbes and the Social Contract Tradition*. Cambridge: Cambridge University Press, 1995.
- Harman, Gilbert. *The Nature of Morality: An Introduction to Ethics*. New York, NY: Oxford University Press, 1977.
- Hibral, Jason. "Animals, Agency, and Class: Writing the History of Animals from Below." *Human Ecology Review* 14, no. 1, 101–112.
- Hilden, Julie. "A Contractarian View of Animal Rights: Insuring Against the Possibility of Being a Non-Human Animal." *Animal Law* 14, no. 5 (2007): 5–28.
- Hobbes, Thomas. *Leviathan: With Selected Variants from the Latin Edition of 1668*. Edited by Edwin Curley. Indianapolis: Hackett Publishing Company, 1994.
- . "Philosophical Rudiments Concerning Government and Society (Translation of De Cive)." In *The English Works of Thomas Hobbes of Malmesbury*, edited by Sir William Molesworth, vol. 2. London, UK: John Bohn, 1839.
- Hume, David. *A Treatise of Human Nature*. Edited by Ernest C. Mossner. London, UK: Penguin Books, 1969.

- Hume, David. *An Inquiry Concerning the Principles of Morals*. Edited by Charles W. Hendel. New York: The Liberal Arts Press, 1957.
- Humphreys, Rebekah. "Contractarianism: On the Incoherence of the Exclusion of Non-Human Beings." *Percipi* 2 (2008): 28–38.
- Jesseph, Douglas. "Hobbes and the Method of Natural Science." In *The Cambridge Companion to Hobbes*, edited by Tom Sorell, 86–107. Cambridge: Cambridge University Press, 1996.
- Kant, Immanuel. *Fundamental Principles of the Metaphysics of Morals*. Translated by T. K. Abbott. Amherst, NY: Prometheus Books, 1988.
- Kymlicka, Will. *Contemporary Political Philosophy: An Introduction*. 2nd ed. Oxford: Oxford University Press, 2002.
- La Barbera, Christopher. *States of Nature: Animality and the Polis*. New York: Peter Lang, 2012.
- Larrère, Catherine, and Raphaël Larrère. "Animal Rearing as a Contract?" *Journal of Agriculture and Environmental Ethics* 12, no. 1 (2000): 51–58.
- Luper, Steven. "Annihilation." In *Life, Death, & Meaning, Second*, edited by David Benatar, 201–222. Toronto: Rowman & Littlefield, 2010.
- Mackie, John L. *Ethics: Inventing Right and Wrong*. New York, NY: Penguin Books, 1977.
- Mill, John Stuart. "Utilitarianism." In *Utilitarianism and On Liberty, Second*, edited by Mary Warnock, 181–235. Malden, MA: Blackwell Publishing, 2003.

- Morris, Christopher W. "Moral standing and rational-choice contractarianism." In *Contractarianism and Rational Choice: Essays on David Gauthier's Morals By Agreement*, edited by Peter Vallentyne, 76–95. Cambridge: Cambridge University Press, 1991.
- . "The Idea of Moral Standing." Chap. 9 in *The Oxford Handbook of Animal Ethics*, edited by Tom L. Beauchamp and R.G. Frey, 255–275. Oxford University Press, 2011.
- Nagel, Thomas. "Hobbes's Concept of Obligation." *Philosophical Review* 68 (1959): 68–83.
- Narveson, Jan. "Animal Rights." *Canadian Journal of Philosophy* 7, no. 1 (1977): 161–178.
- . "Animal Rights Revisited." In *Ethics and Animals*, edited by William H. Williams and Harlan B. Miller, 45–59. Clifton, N.J.: Humana Press, 1983.
- . "On a Case for Animal Rights." *The Monist* 70, no. 1 (1987): 31–48.
- . *Respecting Persons in Theory and Practice: Essays on Moral and Political Philosophy*. Lanham: Rowman & Littlefield, 2002.
- . *The Libertarian Idea*. Peterborough, Ontario: Broadview Press, 2001.
- Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.
- Nussbaum, Martha C. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, Mass.: The Belknap Press: Harvard University Press, 2006.

- Oma, Kristin Armstrong. "Between Trust and Dominion: Social Contracts Between Humans and Animals." *World Archaeology* 42, no. 2 (2010): 175–187.
- Palmer, Clare. *Animal Ethics in Context*. New York: Columbia University Press, 2010.
- "Palmyra: Islamic State locks down ancient city's museum." BBC. 2015. <http://www.bbc.com/news/world-middle-east-32859347>.
- Peters, R. S. *Hobbes*. Harmondsworth: Penguin, 1956.
- Peterson, Martin. *An Introduction to Decision Theory*. Cambridge, UK: Cambridge University Press, 2009.
- Plato. "Euthyphro." In *Plato: Complete Works*, edited by John M. Cooper, translated by G.M.A. Grube, 1–16. Indianapolis: Hackett Publishing Company, 1997.
- . *The Republic of Plato*. Translated by Francis Macdonald Cornford. Oxford: Oxford University Press, 1945.
- Rawls, John. *A Theory of Justice*. Revised Edition. Cambridge, Mass.: The Belknap Press of Harvard University Press, 1999.
- . *Lectures on the History of Moral Philosophy*. Edited by Barbara Herman. Cambridge, Mass.: Harvard University Press, 2000.
- Ray, Debraj. *A Game-Theoretic Perspective on Coalition Formation*. Oxford: Oxford University Press, 2007.
- Regan, Tom. *The Case for Animal Rights*. Berkeley: University of California Press, 1983.

- Rowlands, Mark. "Contractarianism and Animal Rights." *Journal of Applied Philosophy* 14, no. 3 (1997): 235–247.
- Scanlon, Thomas M. *What We Owe to Each Other*. Cambridge, Mass.: Belknap Press of Harvard University Press, 1998.
- Sidgwick, Henry. *The Methods of Ethics*. Seventh. Indianapolis: Hackett Publishing Company, 1981.
- Singer, Peter. *Animal Liberation*. New York: Ecco, 2002.
- . *Practical Ethics*. Third. Cambridge: Cambri, 2011.
- Smith, Kimberly K. "Animals and the Social Contract: A Reply to Nussbaum." *Environmental Ethics* 30 (2008): 195–207.
- Tanner, Julia. "Contractarianism and Secondary Direct Moral Standing for Marginal Human and Animals." *Res Publica* 19 (2013): 141–156.
- Taylor, A. E. "The Ethical Doctrine of Hobbes." In *Hobbes Studies*, edited by Keith Brown, 35–55. Oxford: Blackwell, 1965.
- Taylor, Matthew. "Grounding Animal Rights in Mutual Advantage Contractarianism." *Les ateliers de l'éthique / The Ethics Forum* 9, no. 3 (2014): 184–207. doi:10.7202/1029065ar.
- Thomas, Elizabeth Marshall. *The Hidden Life of Dogs*. Boston: Houghton Mifflin, 1993.

- Tucker, C., and C. MacDonald. "Beastly Contractarianism? A Contractarian Analysis of the Possibility of Animal Rights." *Essays in Philosophy* 5, no. 2. Animal Ethics, Article 31 (2004).
- Vallentyne, Peter. "Contractarianism and the Assumption of Mutual Unconcern." In *Contractarianism and Rational Choice: Essays on David Gauthier's Morals By Agreement*, edited by Peter Vallentyne, 71–75. Cambridge: Cambridge University Press, 1991.
- Vonne Lund, Raymond Anthony, and Helena Röcklinsberg. "The Ethical Contract As a Tool in Organic Animal Husbandry." *Journal of Agriculture and Environmental Ethics* 17 (2004): 23–49.
- Warrender, Howard. *The Political Philosophy of Hobbes*. Oxford: Clarendon Press, 1957.
- Watkins, John W.N. *Hobbes's System of Ideas*. London, UK: Hutchinson, 1965.