The Question of Exclusion in Rawlsian Contractualism

Areti Theofilopoulou

St Cross College

Thesis submitted in partial fulfilment of the requirements for the degree of
DPhil in Philosophy in the Faculty of Philosophy at
the University of Oxford

Hilary 2019

74,227 words

The Question of Exclusion in Rawlsian Contractualism

Areti Theofilopoulou

St Cross College

Thesis submitted in partial fulfilment of the requirements for the degree of
DPhil in Philosophy in the Faculty of Philosophy at
the University of Oxford

Hilary 2019

Abstract

This thesis focuses on what I call the question of exclusion. This question, I argue, is one that poses serious challenges to social contract approaches to justice and political legitimacy. In an intuitive way, the exclusion of some individuals seems to be a corollary of the social contractualist approach, which ascribes justice or legitimacy to a social arrangement insofar as it can be regarded as the product of the (actual – expressed or tacit – or hypothetical) consent of specified parties. Consequently, this approach excludes those whose consent is not required in order to view an arrangement as legitimate or just.

Given the enormous and continuing influence of John Rawls's social contract theory, I take the Rawlsian view as my primary focus, concentrating on the way in which the question of exclusion raises challenges for Rawlsian contractualism. According to Rawls's theory, those who participate in the hypothetical contract and to whom political power ought to be justifiable are *reasonable* and *rational*. They are reasonable in the sense that they understand and are willing to comply with the requirements of justice and they are rational in that they have the capacity to form, revise, and pursue a conception of the good. This gives rise to two categories of exclusion. In the first part of the thesis, I examine the exclusion of those who will develop the two moral powers sufficiently only if those who are represented in the social contract act (or refrain from acting) in certain ways. This category includes children, foetuses, and unreasonable citizens. Intuitively, justice issues in requirements in at least some of these cases, but it is hard to make sense of this on the Rawlsian approach. In the second part of the thesis, I address the question of exclusion regarding those who will *never* develop the two moral powers to the requisite degree, such as individuals with severe cognitive disabilities and non-human animals. Again, intuitively, justice issues in requirements here, but their justification is not clear on the contractualist approach. My conclusion is threefold. First, I argue that the problems the question of exclusion poses cannot be addressed by a single solution. Second, I safeguard Rawlsian contractualism by showing how each problem can be resolved. Third, I offer policy-prescriptions on a number of issues, such as education, abortion, and the criminal justice system.

Acknowledgements

The last 3.5 years taught me more than I could ever hope for. They were full of valuable moments, life lessons, and academic experiences that I could not even imagine when I started this DPhil. I am sincerely grateful to each and every one of my friends and family who contributed to these moments along the way and who helped me shape the way I think about philosophy and life more broadly. I consider myself very lucky to have such incredibly supportive and kind people by my side. It is no overstatement to say that I – as the person that I am –, the way that I think, and the things that I do (including this thesis) would not exist in their current forms had it not been for these people.

I should first express my gratitude to my supervisors, Tom Sinclair and Dominic Wilkinson. The ways that I think and write about philosophy were shaped while I was under their academic care. I know that I am unusually lucky for having had such supportive, kind, and insightful supervisors. I am grateful for all the time that they put into helping me write this thesis, for the valuable, constructive comments that influenced my thesis, papers, and applications tremendously, and, most importantly, for the support that they gave me throughout this DPhil. I would also like to thank Tom Douglas who acted as a mentor even though he didn't have to. I am grateful both for the work that I have done on the project he led, without which pursuing this DPhil would not have been possible, and for the comments and advice that he generously gave me on a number of papers and applications.

For their insightful comments, questions, and helpful conversations, I am grateful to Daniel Butt, who was my Transfer of Status and Confirmation of Status examiner, Hannah Maslen, who co-examined my transfer of status, Jeremy Williams, who generously provided me with lengthy comments on Chapter 3, and the attendants at the Nuffield Workshop on Gerald Gaus's work, the Life, Mind & Death Symposium – Amsterdam-Oxford Consortium, the Nuffield Political Theory Workshop (Michaelmas 2017), the MANCEPT Workshop on Theories of Public Reason (2016 & 2017), the Bucharest – Oxford Workshop in Applied Ethics (2017), the VIII Braga Meetings on Ethics and Political Philosophy, the Oxford Applied Ethics Graduate Discussion Group (Hilary & Michaelmas 2016), the Oxford-Denmark Workshop in Applied Moral Philosophy, and the 13th World Congress – International Association of Bioethics.

I am extremely lucky to have a family that has always encouraged me to live my life according to my own desires and convictions and to be the kind of person that I want to be. This DPhil would not have been possible without the support of my parents, Lina and Takis, who have always encouraged me to form and pursue my own conception of the good and who have supported this pursuit both emotionally and practically. I am also indebted to my sister (and best friend), Alexia, with whom I have shared arguably the majority of both the good and difficult moments of the last three years. For their invaluable support, I would further like to thank Christianna, Thodoris, Mari, Paulos, Anna, Yiayia, and Pappou; I will always be grateful for everything that they have done for me. Finally, I would like to thank my three cousins: Semeli, for all the moments of sheer joy and fun; Daphne, for helping me change my life and for being my best friend; and Mara, for offering me help that I will never forget at the very beginning of the DPhil.

This thesis would certainly not have been possible without the help and presence of my friends. It would be impossible to mention everyone who had a positive impact on me and (by extension) my thesis, so I limit myself to those who had a very direct contribution to this thesis. I am grateful to my good old flrks, Themis Pantazakos for the countless study sessions and for being my flrks; to Evi Bakirtzi, for being the most supportive friend I could ask for; to Angeliki Vourdaki, for consistently making me a better and happier person; to Annette Zimmermann, for discussing *all* the ideas with me, for being my meow friend, and for moments of fun; to Carissa Veliz, for our talks, laughter, and her constant support; to Akanksha Awal, for helping me in ways I could not imagine at a difficult time and for offering me incredible joy; and to Yiannis Tolias, with whom our paths in life have parted, but to whom I will always be grateful for his support during the first three years of the DPhil. For their support, for the moments of fun we shared, and for our academic and non-academic conversations, I am also grateful to Tom Parr, Leonie Schulte, Sabrina Martin, and Diana Popescu.

I would also like to express my gratitude to Antigoni Orfanou, who taught me how to ask the important questions, how to find the even more important answers, and how to be myself and live a happy life.

Finally, I am grateful to the Society for Applied Philosophy and the Sir Richard Stapley Educational Trust, whose funding made writing this thesis possible.

Contents

1. Introduction 9

    1.1 Society as Social Cooperation 9
    1.2 Two Principles of Justice 12
    1.3 The Basic Structure of Society 15
    1.4 Stability and the Turn to Political Liberalism 16
    1.5 The Two Moral Powers 19
    1.6 The Problems of Extension 22
    1.7 The Question of Exclusion 26
        1.7.1 The Question of Exclusion in the Social Contract Tradition 26
        1.7.2 The Question of Exclusion in Rawlsian Contractualism 30
        1.7.3 The Problems Generated by Exclusion 34
        1.7.4 Possible Solutions 37
    1.8 Thesis Outline 41

2. Children 45

    2.1. Introduction 45
    2.2. The Exclusion of Children 48
        2.2.1. Exclusion from the Justificatory Constituency 50
        2.2.2. Exclusion from the Represented Constituency in Justice as Fairness 52
        2.2.3. Exclusion from the Constituency to Which Justice as Fairness Applies 53
    2.3. Avoiding Exclusion 57
        2.3.1. The Interest in Developing the Two Moral Powers 58
        2.3.2. The Interest in Developing the Two Moral Powers and Children's Inclusion 69
    2.4. Constructing Child—Centred Policies 71
        2.4.1. States 73
            2.4.1.1. The Epistemic Aspect of Reasonableness 74
            2.4.1.2. The Epistemic Aspect of Rationality 75
            2.4.1.3. The Motivational Aspect of Reasonableness 78
            2.4.1.4. The Motivational Aspect of Rationality 83
        2.4.2. Parents 84
            2.4.2.1. The Epistemic Aspect of Rationality 84
            2.4.2.2. The Epistemic Aspect of Reasonableness 88
            2.4.2.3. The Motivational Aspect of Rationality 89
            2.4.2.4. The Motivational Aspect of Reasonableness 93
    2.5. Conclusion 96

3. Foetuses 98

    3.1. Introduction 98
    3.2. The Incompleteness Objection 102
        3.2.1. Indeterminacy 102
        3.2.2. Inconclusiveness 107

3.3. The Extreme Permissibility Argument 110
3.3.1. The Objectionable Implications of Public Reason 112
3.4. The Extreme Non-Permissibility Argument 123
3.5. Conclusion 125

4. Unreasonable Persons 127
4.1. Introduction 127
4.2. The Exclusion of Unreasonable Persons 130
4.2.1. Exclusion from the Justificatory Constituency 130
4.2.2. Exclusion from the Constituency of Justice 132
4.3. Why Is Exclusion a Problem? 135
4.4. Including Unreasonable Persons in Rawlsian Contractualism 138
4.5. Implications 142
4.6. Objections 145
4.6.1. The Causes of Unreasonableness 145
4.6.2. The Inequality Objection 148
4.6.3. The Challenge of Indoctrination 149
4.7. Conclusion 153

5. Individuals with Cognitive Disabilities 154
5.1. Introduction 154
5.2. Which Disabilities Raise the Question of Exclusion? 158
5.3. Potential Solutions 163
5.3.1. *Frontiers of Justice* 163
5.3.2. Personhood as Minimal Reasonableness and Rationality 169
5.3.3. Reciprocity and Relational Contributions 173
5.4. Resolving the Problem of Exclusion 177
5.4.1. The Interest in Developing the Two Moral Powers 178
5.4.2. Parental, State, and Citizens' Duties and Opacity Respect 181
5.5. Conclusion 193

6. Non-Human Animals 194
6.1. Introduction 194
6.2. Potential Solutions 202
6.2.1. The Intuitive Equality Argument 202
6.2.2. Human-Animal Relationships 205
6.2.3. Altruistic Contractualism 208
6.2.4. Appealing to an Overlapping Consensus 210
6.2.5. Capabilities and Species Membership 211
6.2.6. An Intuitive Argument for Trusteeship 214
6.3. An Argument from Compassion 215
6.4. Conclusion 222

7. Conclusion 224

Chapter 1

Introduction

John Rawls's influence on Western contemporary political philosophy is undeniable. Ever since the publication of *A Theory of Justice* in 1971, political philosophy has to a great extent engaged with Rawls's work by recasting, extending, or criticising his elaborate arguments. It is partly for this reason that my thesis focuses on Rawls's work. The purpose of this project is to discuss what I will call the question of exclusion. This question, I will argue, is one that poses serious challenges to all social contract approaches to justice and political legitimacy, of which Rawls's theory is but one. Given the significance of Rawls's work in contemporary political philosophy and because I take his theory to be the strongest interpretation of social contractualism, I focus on the way in which the question of exclusion poses a challenge for Rawlsian contractualism. In order to understand this challenge, we need a clear account of the main concepts and ideas in Rawls's theory of justice and legitimacy as this was developed in *A Theory of Justice*, *Political Liberalism*, and *Justice as Fairness: A Restatement*. That is what I provide in the first part of this introduction. In the second half, I set out the problem of exclusion.

1.1 Society as Social Cooperation

Following the social contract tradition, Rawls conceives of society as social cooperation that has the primary purpose of mutual advantage. The idea is that in the absence of the social cooperation that a state secures for its citizens through a system of fair rules and

regulations, individuals would be worse off. As a result, against the baseline of non-cooperation, participating in social cooperation can be advantageous for all. Importantly, the idea of mutual advantage in social cooperation is compatible with different views on human nature and does not require, therefore, accepting a specific account of human nature. For example, those agreeing with Hobbes that the state of nature would be a 'war of all against all' and 'the life of man, solitary, poor, nasty, brutish and short', will clearly view social cooperation as advantageous for all those who participate in it. Similarly, those who disagree with the Hobbesian approach and endorse an account of human nature that stresses the inherently good and social nature of humans will view social cooperation as a relation that expresses and potentially reinforces human nature and, thus, as a relation from which we all benefit. And the idea of mutual advantage in social cooperation is also consistent with the view that human beings are neither completely self-interested nor fully altruistic, but mutually disinterested; for even if we endorse that view, we can still see social cooperation as mutually beneficial, given that it enables us to avoid the tragedy of the commons and to have access to the benefits of both trade and the division of labour. Thus, we see that the social contractualist idea that human interests are furthered by our participation in social cooperation is compatible with the main accounts of human nature. As will become clearer in my discussion of Rawls's *political* liberalism, the fact that we can view mutual advantage as the purpose of social cooperation without endorsing a specific account of human nature or a specific view about which ways of living are good renders this conception of society consistent with the broader aims of Rawlsian contractualism.

Although social cooperation is advantageous to all those who participate in it compared to the baseline of non-cooperation, the benefits and costs that it generates for each person are

determined by the rules that constitute the cooperative system. The task of a theory of justice is to determine the *fair* distribution of the benefits and burdens of social cooperation, which all real societies should strive to materialise.

Before turning to Rawls's response to the question of fairness in this distribution, it is important to note that justice is not expected to be feasible or desirable in all possible worlds. Rather, because justice is linked to social cooperation and mutual advantage, it becomes relevant only when 'human cooperation is both possible and necessary' (Rawls 1971, 126). This is captured by the Humean idea of the circumstances of justice, which highlights that justice is made necessary by both human behaviour and the provisions of nature. In the words of Hume,

> 'tis only from the selfishness and confin'd generosity of men, along with the scanty provision nature has made for his wants, that justice derives its origin (Hume 2000, 495).

The circumstances of justice are thus a combination of objective and subjective factors (Nussbaum 2006, 27). The objective circumstances include conditions such as that the parties live 'at the same time on a definite geographical territory'; that they possess similar physical and mental powers; that they are vulnerable; and that they live under conditions of 'moderate scarcity' – that is, that the earth's resources are neither so abundant as to render cooperation unnecessary nor so scarce as to make cooperation impossible (Rawls 1971, 126-127). As for the subjective circumstances of the parties, the more specific requirement is that they possess 'similar needs and interests, or at least complementary interests', and that they have distinctive conceptions of the good life (Nussbaum 2006, 27). Although the subjective circumstances of justice include facts about the parties' needs, interests, and conceptions of the good, we need not endorse a controversial account of the *specific* needs,

interests, and conceptions of the good the parties would have in a pre-social state of nature. Rather, given that our needs, interests, and conceptions of the good are influenced by environmental factors, and given that Rawlsian contractualism aims to identify the right principles of justice for *us*, we may suppose that the needs, interests, and conceptions of the good of the parties are the ones that we have (Munoz-Dardé and Sinclair, forthcoming). This assumption would satisfy the criteria set by the subjective circumstances of justice without requiring us to implicitly endorse a controversial account of what humans are like in the state of nature.

1.2 Two Principles of Justice

Following the social contract tradition, Rawls views a (hypothetical) agreement between all members of society as the source of principles of justice and legitimacy. However, his theory departs from traditional accounts in two crucial ways. First, as we have seen, his theory does not rely on a specific account of human nature or of what makes life valuable. One of the virtues of his theory of justice is that it is consistent with a number of different views on these issues.

Second, although Rawls's aim is to provide an example of 'pure procedural justice' in his account of the social contract, whereby the principles of justice are correct because they are generated from the right procedures, his construction of the conditions of that agreement imports certain moral elements in the procedure. These moral elements reflect our considered judgments about which factors are morally relevant or arbitrary in the determination of the distribution of the benefits and burdens of social cooperation. The main idea is that any characteristics that do not affect one's capacity to participate in social cooperation are morally irrelevant; and fairness requires that morally irrelevant

characteristics do not affect the distribution of benefits and burdens in society. For instance, individuals' gender is utterly irrelevant to their capacity to participate in social cooperation and, by extension, to the fair distribution of the benefits and burdens of that cooperation. Traits such as species membership, on the other hand, are not on a par with traits such as gender, given that they do affect one's capacity to participate in social cooperation. Thus, instead of merely asking what principles the parties would agree to in the state of nature – or the Rawlsian equivalent, the Original Position – the parties are placed under the 'Veil of Ignorance', which makes them ignorant of some of their characteristics, such as gender, but not others, such as species membership. In this way, the parties are not aware of the characteristics that ought to be viewed as morally arbitrary, precisely because these characteristics do not affect their capacity to participate in social cooperation. The 'Veil of Ignorance', then, models our assumptions about the inappropriateness of giving a role to the parties' asymmetrical power in their deliberation. As individuals are bound to have different natural and social circumstances and as different natural talents and abilities are valued differently in different societies, in the absence of the Veil of Ignorance, some parties would possess considerably more bargaining power than others. By contrast, the Veil of Ignorance expresses the idea of justice as fairness in the sense of moral impartiality with regards to the characteristics that are appropriately viewed as irrelevant to the parties' capacity to participate in social cooperation. By not being aware of their own relative position in society, the parties are prevented from making decisions based on bargaining advantages that they have no special claim to have in determining the terms of cooperation.

The parties' choice occurs under the further assumption that their society is closed and self-sufficient: the parties must imagine, that is, that there is no possibility of migration and no need for international trade. These assumptions are made because they ensure that the

parties' reasoning cannot be affected by the prospect of emigrating to another country where they could live under more advantageous terms (Rawls 2005, 41). In the absence of such prospects, the parties are motivated to choose the principles that it is most rational for them to choose for their own society, which, given the design of the Original Position, are also the fairest principles they could choose. Although there are important questions about global justice that arise from these assumptions, it would be impossible to address them here; these questions are pressing but lie beyond the scope of this thesis.

Under these circumstances, Rawls stipulates, the parties are presented with the main conceptions of justice – namely, utilitarianism, perfectionism, and intuitionism, and would compare them to the set of principles given by Justice as Fairness (Rawls 2001, 97). Under the conditions of the Original Position, he argues, the parties would rationally choose the principles of Justice as Fairness. One way of understanding how the parties would reach this decision is through the maximin rationale, although, as Rawls stresses, this is simply one among other heuristic devices that can be used to enable us to realise what our basic interests are (Rawls 2001, 99). Briefly, according to the maximin rationale, the parties would choose the principles that would maximise the benefits for the least advantaged members of society, out of concern that they might turn out to be among these members. This rationale favours the principles of Justice as Fairness:

> First Principle [Liberty Principle]: Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all;
>
> Second Principle: Social and economic inequalities are to satisfy two conditions:
>
> a. [Fair Equality of Opportunity] They are to be attached to offices and positions open to all under conditions of *fair equality of opportunity*;

> b. [Difference Principle] They are to be to the greatest benefit of the least-advantaged members of society (Rawls 2001, 42–43).

1.3 The Basic Structure of Society

These principles are not, however, to be applied to any interaction, institution, or individual in society. Rather, their purpose is to ensure that the main social and political institutions that are needed for the purpose of social cooperation distribute the benefits and burdens of that cooperation justly. These institutions, called the basic structure of society, are described by Rawls in the following passage

> For us, the primary subject of justice is the basic structure of society, or more exactly, the way in which the major social institutions distribute rights and duties and determine the division of advantages from social cooperation. By major institutions I understand the political constitution and the principal economic and social arrangements. Thus the legal protection of freedom of thought and liberty of conscience, competitive markets, private property in the means of production, and the monogamous family are examples of major social institutions. Taken together as one scheme, the major institutions define men's rights and duties and influence their life prospects, what they can expect to be and how well they can hope to do. The basic structure is the primary subject of justice because its effects are so profound and present from the start (Rawls 1971, 7).

This may appear somewhat vague, for it does not specify precisely which institutions are part of the basic structure. It seems, however, that this vagueness is intended to highlight the fact that the content of the basic structure is socially contingent, as different institutions may exist and have this pervasive role in different societies. In each case, the institutions that form a society's basic structure will meet two necessary and sufficient conditions. The first condition is that they will have profound effects on citizens' lives, as they will form not only their opportunities and prospects in life, but their character, aspirations, capacities, and attitudes as well. Moreover, the reason why these institutions have these effects 'cannot

possibly be justified by an appeal to the notions of merit or desert' (Rawls 1971, 7). In any case, whatever the precise form of the basic structure, it should be noted that the principles of justice are designed to apply to the basic structure as a whole, and not to each institution individually (Munoz-Dardé and Sinclair, forthcoming).

1.4 Stability and the Turn to Political Liberalism

According to Rawlsian contractualism, a well-ordered society should be stable over time. This implies that its basic structure should have the capacity to remain just with the passage of time. As long as its citizens are sufficiently reasonable, this would give them sufficient reasons to support their society's institutions. Moreover, the well-ordered society should be stable *for the right reasons*. That is, the reasons for which its citizens should accept and support that society's institutions should not be ones of fear or necessity that simply ensure society's durability; in that case, that society would endure only as a *modus vivendi* at best (Rawls 2005, xxxvii). Rather, the well-ordered society should generate support for its institutions on the basis of its citizens' commitments, values, and conceptions of the good.

The idea of stability for the right reasons gives rise to the liberal principle of legitimacy, according to which

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

A society is therefore legitimate when it can be stable for the right reasons. For if it is stable for the right reasons, by definition, it provides its citizens with good reasons for its acceptability (Munoz-Dardé and Sinclair, forthcoming). We thus see that the question of

stability is, at its heart, a question of legitimacy; it asks under what conditions a well-ordered society can be stable *for the right reasons* and, by extension, legitimate.

One potential challenge to stability for the right reasons, however, is the existence of reasonable disagreement on the good, given that individuals hold different comprehensive doctrines.[1] This kind of disagreement is seen as the inevitable outcome of the 'burdens of judgment' under liberal institutions (Rawls 2005, 57). For instance, even if we agree on the values that we should take into consideration when discussing issues of justice, we often disagree on the weight that we attach to these considerations. This is because our experiences influence our assessment of different facts and values to such an extent that even similar individuals with similar backgrounds tend to have experiences that give rise to disagreement (Rawls 2005, 57). For these reasons, reasonable disagreement is embraced as the natural outcome 'of the exercise of human reason within the framework of the free institutions of a constitutional democratic regime' (Rawls 2005, xvi). For if disagreement exists in virtue of the very way in which humans freely reason, a society without it can only exist if it suppresses human reason itself.

Yet even though disagreement is not to be lamented, it gives rise to difficulties. This is because, according to the liberal principle of legitimacy and the concept of stability *for the right reasons*, constitutional essentials and issues of basic justice must be justifiable to reasonable citizens. If, however, reasonable citizens disagree on the comprehensive

[1] Rawls defines comprehensive doctrines in the following way: "a moral conception is general if it applies to a wider range of subjects, and in the limit to all subjects universally. It is comprehensive when it includes conceptions of what is of value in human life, and ideals of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole. A conception is fully comprehensive if it covers all recognized values and virtues within one rather precisely articulated system; whereas a conception is only partially comprehensive when it comprises a number of, but by no means all, non-political values and virtues and is rather loosely articulated. Many religious and philosophical doctrines aspire to be both general and comprehensive' (Rawls 2005, 13).

doctrines that they endorse, then it seems that a society's exercise of political power on these issues must be consistent with each of these reasonable comprehensive doctrines affirmed by its citizens. Consequently, it should not implicitly or explicitly favour a particular comprehensive doctrine.

We thus reach two requirements. First, the theory of justice that determines the distribution of benefits and burdens in a well-ordered society must be political rather than comprehensive; that is, its principles, its theoretical justification, and its account of stability should not rely on any controversial comprehensive doctrines. Second, the well-ordered society's inquiry must be free and public: the actions of its public officials and its citizens, when acting qua citizens, must be justified in terms of political, and not comprehensive, values (Rawls 2005, 224). This second requirement expresses the idea of public reason. In accordance with the wide view of public reason, which Rawls endorsed right after the publication of the first edition of *Political Liberalism*, public officials and citizens can introduce comprehensive doctrines in the 'public political discussion at any time, provided that in due course proper political reasons – and not reasons given solely by comprehensive doctrines – are presented that are sufficient to support whatever the comprehensive doctrines introduced are said to support' (Rawls 2005, 462).

The question then becomes whether Rawls's theory of justice can be stable for the right reasons without appealing to a comprehensive account of the good; or, in other words, whether a state that follows Rawls's principles of justice can also be legitimate. An affirmative answer requires that the theory be recast as a political, 'freestanding' conception, whose content, including its account of stability, would not require the acceptance of specific comprehensive doctrines. As the theory's principles and their

justification were already cast in terms that can be viewed as political in *A Theory of Justice*, the main change that was needed was in the theory's account of stability. This is because the arguments that Rawls presented in favour of the well-ordered society's stability in his early work are typically seen as appealing to comprehensive views.[2]

In order to show why the well-ordered society of justice as fairness would be stable for the right reasons, Rawls appeals to the concept of 'overlapping consensus', which states that the theory can seem acceptable to citizens who hold different, sometimes conflicting, reasonable comprehensive doctrines (Rawls 2005, 65; 134-137). This is because the very way in which a political conception of justice is formulated ensures that it is justifiable to all reasonable citizens, despite their reasonable disagreement on issues related to comprehensive doctrines. As Rawls stresses in *Political Liberalism*, being the object of an overlapping consensus in a well-ordered society is something that other theories of justice can achieve, as long as they are political rather than comprehensive, and respect key liberal values. Justice as fairness should therefore be viewed as one conception in a family of liberal conceptions of justice that can be stable for the right reasons.

1.5 The Two Moral Powers

Given that political liberalism (and by extension, justice as fairness) cannot rely on a metaphysical account of personhood, it needs a political account of personhood in order to describe the nature of persons qua citizens of the well-ordered society. According to this account, persons are viewed as possessing two key moral powers, developed at least to the minimum degree that enables them to participate in social cooperation. This conception of

[2] The precise details of Rawls's account of stability in *A Theory of Justice* lie beyond the scope of this thesis. For a detailed analysis, see Rawls 1971, §78, §79, §86; Munoz-Dardé and Sinclair, forthcoming; Freeman 2002; Weithman 2011.

personhood is a political one precisely because it is derived from the capacity for social cooperation (Rawls 2005, 370).

The first moral power, reasonableness, is the capacity to understand the requirements of justice and the willingness to comply with those requirements, on the assumption that others are prepared to do so as well. In Rawlsian contractualism, reasonable persons further acknowledge the burdens of judgment and accept, for that reason, that they ought to comply with the requirements imposed by public reason.[3]

Rationality, on the other hand, which is the second moral power, is citizens' capacity to form, revise, and pursue a conception of the good. Such conceptions may include ideas about what makes life worth living, views on human character and virtue, as well as other comprehensive doctrines. If persons lacked rationality, they would lack a crucial condition for freedom; for it is rationality that endows us with the capacity to lead lives that can be viewed as authored, or in some way shaped by us. Relatedly, in the absence of rationality, individuals would not have the capacity to identify and pursue what would count as rational advantage *for them* when they are meant to be cooperating with others with the purpose of *mutual* advantage (Rawls 2005, 300). That is, without rationality, individuals would not have a conception of the good that they have good reason to pursue by participating in social cooperation.

Reasonable and rational persons thus have the capacity to form, revise, and pursue *reasonable* conceptions of the good – that is, conceptions that are consistent with the

[3] This second part of reasonableness is rejected by a number of Rawlsian political liberals who hold that accepting the burdens of judgment is not a necessary condition for reasonableness (Quong 2011, 158; Klosko 2000, 20—4; Wenar 1995, 41—8).

requirements of justice. In this sense, the reasonable ought to subordinate the rational (Rawls 2005, 367). The way in which this is achieved in Justice as Fairness is through a division of labour between the design of the Original Position and the specifications of the parties that participate in the hypothetical agreement. On the one hand, the parties are modelled in accordance with the idea of rationality: they are mutually disinterested individuals who seek to maximise their rational advantage. This account of rationality does not require a commitment to egoism; there is no denial that the people represented by the parties can be genuinely altruistic. It merely implies that the parties that represent these individuals seek to protect their interests, whatever these may be, rather than subordinating them to those of others. On the other hand, the design of the Original Position models the idea of reasonableness: as we have seen, by placing the parties under the Veil of Ignorance, it is ensured that their mutually disinterested decisions do not reflect unfair bias or partiality (Rawls 2005, 370).

As I have mentioned, the two ideas form the political conception of personhood because they are viewed as necessary and sufficient for a citizen's participation in social cooperation. Together, the two ideas further ground all participants' freedom and equality. In virtue of possessing the two capacities, individuals are free (Rawls 2005, 19). The two moral powers are linked to freedom in two related ways. First, in order to be, and view themselves as, independent and 'self-authenticating sources of valid claims', individuals must be reasonable and rational (Rawls 2005, 32). To see why, consider, for instance, the case of someone who is rational but not reasonable. Despite having the capacity to form, revise and pursue her conception of the good, the self-authenticating claims that arise from that person's conception of the good would not be valid, given that they would not be consistent with justice. Similarly, if one were reasonable but not rational, one would lack

the capacity to be a *self-authenticating source* of a claim regarding the way in which she should lead her life. Second, possessing the two moral powers ensures that citizens are in a position to view themselves as responsible for the ends that they choose to have (Rawls 2005, 33).[4] Assuming that being an agent who can be held morally responsible for her decisions is part of what it means to be a free person who can make free choices, possessing the two moral powers is a necessary condition for freedom. This explains why, for instance, infants cannot be viewed as free persons while they are infants: for they lack the capacity to be responsible self-authenticating sources of valid claims. The possession of these capacities to the sufficient minimum degree further implies that individuals are equal (Rawls 2005, 19). The two moral powers specify the basis of equality because all that matters for the purposes of social cooperation is that citizens are sufficiently capable of participating in it according to the rules and regulations that are specified by justice.

1.6 The Problems of Extension

Having set out the basic framework of Rawls's theory of justice and legitimacy, I now turn to what I will call the question of exclusion. To see how this question arises, recall that Rawls's theory relies on a set of assumptions that are not true of our societies here and now. The assumptions that are relevant here are that the members of society are reasonable and rational, and that the society is closed and self-sufficient. These assumptions are not meant to track characteristics that society does or should have but serve to stress the significance of social cooperation, as I explained above.

[4] It should be clear that this conception of freedom does not imply that individuals are *metaphysically* prior to or responsible for their ends.

The reliance on these assumptions about citizens and society automatically creates problems for the discussion of a number of pressing moral and political issues. Rawls terms these issues 'the problems of extension' to reflect his conjecture that we can extend the conclusions we reach about justice and legitimacy to address cases that are initially left aside due to their incompatibility with the assumptions made in the context of ideal theory (Rawls 2005, 245). According to Rawls,

> there are at least four such problems. One is extending justice to cover our duties to future generations (under which falls the problem of just savings). Another is the problem of extending it to the concepts and principles that apply to international law and political relations between peoples – the traditional *jus gentium*. A third problem of extension is that of setting out the principles of normal health care; and finally, we may ask whether justice can be extended to our relations to animals and the order of nature (Rawls 2005, 244-5).

As far as the first three problems are concerned, Rawls asserts that the relevant extension of political liberalism can be achieved 'forward to other generations, outward to other societies, and inward to those requiring normal health care' (Rawls 2005, 245). For example, instead of including representatives from different generations in the Original Position, Rawls suggests that we identify the requirements of intergenerational justice in the following way. We can imagine that the parties would 'be required to agree to a savings principle subject to the further condition that they must want all *previous* generations to have followed it' (Rawls 2005, 274). In this way, it is envisaged that political liberalism does not require abandoning its core idealising assumptions but can in fact add a new constraint in order to extend the idea of justifiability to future generations.

Similarly, regarding the second problem, instead of revising the Original Position to include representatives of all *individuals* (in the world) rather than of all *citizens* (of a single, closed

society), Rawls argues that his theory of domestic justice can be extended to other peoples in a different way: his theory of global justice begins by accepting his conclusions about the requirements of domestic justice and goes on to ask how societies structured according to the requirements of a conception from the family of liberal views on justice ought to approach foreign policy (Rawls 1999, 10).

As an example of how the third problem might be resolved, consider that the prospect of not receiving treatment in cases of illness or injury would not be justifiable to reasonable and rational citizens. This is because in a contractualist scenario such as the Original Position, the contractors would be aware of the statistical possibilities of illness and injury and they would ensure that they safeguard their future selves against such possibilities. In Justice as Fairness in particular this reasoning seems to be the same maximin rationale that can lead the parties to accept the difference principle. Consequently, in these cases, the primary aim of healthcare must be to restore individuals 'by health care so that once again they are fully cooperating members of society' (Rawls 2005, 184).

As for the fourth problem, the issue of animals and the order of nature, Rawls says that this must be addressed by considering the political values at play, such as the benefits of preserving the natural order to human health (Rawls 2005, 245). This case is clearly treated differently from the first three problems of extension because, unlike the subjects of the first three cases, animals are neither currently nor potentially reasonable and rational moral agents and cannot participate in social cooperation as cooperators.[5] Animals and nature are therefore 'seen as subject to our use and wont' (Rawls 2005, 245).

[5] It should be noted here that the assumption that non-human animals cannot participate in social cooperation has been disputed. I rely on this assumption at this point because it offers the mainstream interpretation of what it means to participate in social cooperation but I do address the objection in Chapter 6.

Although Rawls did not offer the above solutions as a fully developed solution to the problems of extension, he presented them as a way of justifying his conjecture that these problems can be addressed by appropriately extending the principles and conclusions of Rawlsian contractualism.

In the remainder of this chapter, I cast doubt on this hopeful conjecture. I first examine the ways in which social contract theories in general face a problem having to do with the way in which they exclude certain groups from the constituency that designs the principles of justice and from those to whom these principles apply. Subsequently, I show how the problem applies to Rawlsian contractualism, and argue that Rawls's 'problems of extension' are really instances of the general question of exclusion. According to my account, the source of these problems lies in the assumptions that underlie Rawls's views on moral personality.[6] Briefly, the idea is that by assuming that the parties in the Original Position are reasonable and rational, and that the two moral powers form the basis of individuals' freedom and equality, two distinct problems of exclusion arise. One is the problem of exclusion regarding those who will never gain the two moral powers to the degree that is required in order to count as a person. The second problem concerns those who can develop the two moral powers to the appropriate degree only if current persons act

[6] It follows that the third problem of extension that Rawls identified, *jus gentium*, is excluded from my discussion, as it arises due to the role of social cooperation in Rawls's theory, and not due to its account of political personhood. I should stress that my discussion of these problems is not exhaustive and is limited to the problems that arise due to Rawls's account of moral personality. The question of what should happen with citizens of other states is a very important one, but one of a different nature, because it occurs due to focusing on the role of actual social cooperation instead of the capacity for social cooperation. There is already a rich literature questioning this role in the context of global justice, but it is interesting to note that even if we accept it, non-citizens are not excluded in the way that non-persons are. This is because in our world as it is, social cooperation occurs on a global level. Therefore, given that non-citizens participate in the social cooperation that occurs in other societies, it seems straightforward that they are not excluded from the constituency of justice.

(or refrain from acting) in certain ways. Finally, I explain why these problems may be particularly threatening to Rawlsian contractualism and consider some possible solutions that involve revising the theory as a whole. I argue in response that the problems must be addressed individually, in distinctive ways. The remainder of the thesis does just that.

1.7 The Question of Exclusion

1.7.1 The Question of Exclusion in the Social Contract Tradition

In an intuitive way, the exclusion of some individuals, which is what causes the 'problems of extension', seems to be a corollary of the social contractualist approach. That approach ascribes justice or legitimacy to a social arrangement insofar as it can be regarded as the product of the (actual – expressed or tacit – or hypothetical) consent of specified parties. Consequently, the approach excludes those whose consent is not required in order to view an arrangement as legitimate or just.

Broadly, there are three approaches to the social contract, all of which imply the exclusion of some individuals. First, we might hold that a social contract actually occurred at some point in time, as an explanation of how societies were first formed and have developed since then, and tie justice and legitimacy to the satisfaction of the terms of this contract. In order to determine whether justice and legitimacy are realised in any situation today, we ask whether our current situation accords with the terms that those party to the social contract agreed, or were most likely to have agreed. In this scenario, we may presume that the terms of the contract would have been in the interests of those agreeing to it. Consequently, the interests of those who were not among the contractors, either because they were not members of the relevant community (such as contemporary outsiders or future generations) or because they were not considered sufficiently competent at the time to be contractors

(such as children), risk not being represented in the decision-making process. Most crucially, in cases where the interests of those excluded clash with the interests of those agreeing to the contract, it seems that this social contractualist approach will inevitably sacrifice the former for the sake of the latter. As a result, according to this approach, not only would current persons not have any duties towards those who cannot participate in social cooperation, they might well have duties to favour those with whom they cooperate in ways that actively set back the interests of non-participants.

The second approach instructs us to imagine a hypothetical social contract in which individuals, as they are now, agree on terms of cooperation, which are taken to be just or legitimate in virtue of that agreement. According to this approach, we need not consider the social contract as a historical event. When we assess the justice or legitimacy of an arrangement, we need not ask what the interests of the inhabitants of a historical state of nature would have been. Rather, we imagine the contract as an agreement that could occur at any point in time, in some cases under idealised conditions, between existing individuals who would base their decisions on the interests that they currently have. This approach is similarly prone to problems of exclusion, however. First, those who are included may not be included on fair terms. This is because the individuals taking part in the hypothetical social contract are as we are, with their own conceptions of the good, characteristics, and abilities, and with different amounts of (bargaining) power. Yet the characteristics that typically endow us with more power here and now, such as social class and gender, are factors that, as we have seen, are morally arbitrary when it comes to determining the fair distribution of the benefits and burdens of social cooperation. Thus, we see that, on this approach, even those who seem to be included may face questions of discrimination and disadvantage that may effectively amount to partial exclusion. Second, and more pertinently

to the purposes of this thesis, some individuals may be entirely excluded from the constituencies of justice and legitimacy because they are not members of the imagined contracting group, such as non-citizens, future people, foetuses, and non-human animals.

According to the third approach, which most contemporary social contract theorists endorse in some form, in order to identify the requirements of justice and legitimacy, we are to imagine a hypothetical contract that would be agreed to by idealised individuals under idealised conditions. In this case, the boundaries of inclusion depend on the terms of idealisation. Suppose, for example, that we are called to imagine an Original Position in which all parties have a minimum IQ of 90, are mutually disinterested and motivated by considerations of their own welfare. Given these assumptions, we can predict that the principles with which these parties will come up will be shaped in accordance with interests that may not match those of real individuals to whom the resulting requirements of justice and legitimacy will apply. Three forms of exclusion may arise in particular. First, if only persons with an IQ over 90 are included in the justificatory constituency, it is possible that the conditions for their state's legitimacy would not include any requirements of justifiability to those with an IQ under 90. Relatedly, this implies a potential exclusion from the realm of justice: assuming that the goal of the social contract is the contractors' mutual advantage, the decisions of these individuals regarding the principles that will regulate their cooperation will presumably best promote the interests of everyone who happens to be in their position. Consequently, even if the principles they agree to apply to those who have an IQ below 90, they may do so in a way that does not promote these persons' interests. Third, we may presume that, in some cases, the principles will not even apply to those who are excluded. For if the contractors agree to principles that will determine the most just distribution of the benefits and burdens of *their own* cooperation, it seems likely that they

will not include those who do not participate in their scheme of cooperation in those to whom the benefits and burdens of that cooperation are distributed. Returning to our example, for instance, individuals with an IQ below 90 may not be granted the protections of the difference principle.

Of course, the IQ example should only be taken as an indication of how exclusion can arise within the third social contract approach, and not as criticism of any particular theory. The general point here is that, given that the theoretical construction of the social contract relies on certain assumptions about who is represented in the social contract situation, certain groups will inevitably be excluded. These assumptions might range over the parties' mental, physical or other properties to temporal location or species membership. Thus, those who do not share the relevant properties are excluded both in the sense that they are not represented in the constituencies of legitimacy and justice, and in the sense that they are excluded from those to whom the principles of justice apply.

At this point, someone might object that *this* exclusion is not particularly objectionable because, even if those excluded are excluded from the realm of *justice* both in the sense that they are not represented in the relevant constituency and in the sense that the principles of justice do not apply to them, they might still be included in the constituency of *morality*. This would imply that there are certain moral duties owed to them. For example, it might be argued that those included in the political community have certain duties of beneficence towards those excluded. However, even in cases where this is true, it may still be objected that there is not sufficient justification to ground the enforceability of these duties. In order to assess the validity of this objection, one must examine the normative foundations of these duties, and their consistency with the commitments and assumptions of the social contract

theory. While it might be true that some moral theories, such as utilitarianism, can escape the charge if they focus on properties, such as sentience, that a greater number of groups share, other theories, such as social contractualism, cannot. This is because the latter determines the content of justice and legitimacy by appealing to the idea of the social contract, and the related idea of mutual advantage as the purpose of social cooperation. According to these ideas, a principle is just depending on whether those who participate in a social contract did or would or could accept it as a principle that determines the distribution of the benefits and burdens of their cooperation.

1.7.2 The Question of Exclusion in Rawlsian Contractualism

As a theory following the social contract approach, Rawlsian contractualism is vulnerable to the problem of exclusion. Having presented the intuitive way in which social contract theories give rise to the question of exclusion, in this section I explore the particular kinds of exclusion that arise within Rawls's theory.

As we have seen, Rawls's account of contractualism is a *political* one: instead of claiming that contractualism fully captures the requirements of morality, this account of contractualism refrains from taking a position on broader moral questions. Rather, it focuses only on the *political* issue of justice in the distribution of the benefits and burdens of social cooperation, while acknowledging that any reasonable moral and metaethical theory might be true.

As the previous section shows, the definition and justification of a contractualist theory's idealised conditions is of great importance, as it sets the boundaries of both the content of justice and the constituency to whom it applies. In Rawlsian contractualism, the

constituencies of justice and legitimacy only include reasonable and rational persons.[7] This means that, as we have seen in the liberal principle of legitimacy, only reasonable and rational persons are included in the justificatory constituency of those to whom the exercise of political power ought to be justifiable in order to be legitimate; relatedly, this also means that only reasonable and rational persons are represented in the constituency that decides on the content of justice; and finally, it implies that there is reason to think that only reasonable and rational persons are included in those to whom the principles of justice apply.[8] As we have seen, the reason why only these persons are included in the constituencies of justice and legitimacy is that only they are sufficiently capable of participating in society construed as social cooperation with the purpose of mutual advantage.

Given the political conception of the person, there are three broad groups to which the question of exclusion applies:

i. Rational but not reasonable individuals
ii. Reasonable but not rational individuals
iii. Those who are neither reasonable nor rational

First, there are those who are rational but unreasonable, either because they lack the capacity to *understand* the requirements of justice or because they lack the willingness to *comply*

[7] Recall that, in this context, being reasonable means that one has the capacity to understand and comply with the requirements of justice, while being rational means that one has the capacity to form, revise, and pursue a conception of the good.
[8] Of course, with the exception of a few principles, such as the natural duty of justice and the principle of fairness, the principles of justice apply to the basic structure of society. By claiming that the principles do not apply to certain individuals, I hereafter mean that the rights protected by the two principles of justice are not assigned to these individuals *and* that the principles that do generally apply to individuals do not apply to these specific individuals.

with these requirements. For example, those who hold views or act in ways that are fundamentally incompatible with the requirements of justice would fall into this category.[9]

The second category of exclusion concerns those who are, conversely, reasonable but not rational.[10] An example of how this might be possible is found in the case of individuals with severe mental health issues, such as bipolar disorder or debilitating depression, which prevent them from forming, revising, or pursuing their conception of the good, without interfering with their capacity to understand and act in accordance with the requirements of justice. Of course, irrationality in this sense comes in degrees and it seems unlikely that any person that falls within this category lacks the capacity completely. However, there might be individuals who fall under the threshold because mental illness or incapacity prevents them either from being able to form a conception of the good, or from having the desire and the capacity to pursue their ends. Yet it seems that in all of these cases, these individuals have once been sufficiently reasonable and rational or go through periods during which they are above the threshold of rationality and can sustain these periods with the appropriate medication. For that reason, this type of exclusion is resolved like standard cases of illness or accident: upon realising that their rationality might diminish during the course of their lives, the parties in the Original Position would choose to insure against such possibilities.

Lastly, and most straightforwardly, there is the exclusion of those who are neither reasonable nor rational. This category includes non-human animals, individuals with severe cognitive disabilities, foetuses, infants, and future generations.

[9] Racist views, for example, are unreasonable because they demonstrate a lack of either willingness or the capacity to understand and potentially follow the requirements of justice, which include the commitment to all citizens' freedom and equality. Insofar as it relies on the key premise that individuals are not equal, racism is a fundamentally unreasonable doctrine.

[10] It is important to keep here in mind that 'rational' is not used in the standard, colloquial way, but as a term that identifies individuals' capacity to form, revise, and pursue a conception of the good.

There is a further important distinction to be made within each category. On the one hand, some individuals falling under the main three categories have the *potential* to develop the two moral powers, but only if those who currently count as persons act (or refrain from acting) in certain ways. A clear example here is found in the case of future generations, members of which will come into existence and develop the two moral powers only if current persons decide to procreate and nurture their children. On the other hand, there are some who will *never* develop the two moral powers that are necessary for the status of political personhood to a sufficient degree. To some extent, precisely who falls within this category depends on medical discoveries and technologies and is therefore subject to change. Clearly, there are cases of treatable impairments or diseases that once seemed permanent. Yet even in the face of change and uncertainty, it seems most likely that there will always be some individuals and non-human animals that will fall into this category.

Before I turn to examine the problems that exclusion poses to Rawlsian contractualism and the way in which the particular cases mentioned above relate to these problems, one clarification is in order. It is clear that different individuals within each excluded group might have the potential to develop the two moral powers, while others will never develop them. For instance, some unreasonable persons might develop the capacity for a sense of justice, while others, such as psychopaths, are less likely to develop reasonableness. Similarly, as we will see in the final chapter, some individuals with cognitive disabilities might end up developing the two moral powers, while others – with severe or profound disabilities – will never develop them. For reasons of clarity, I will assume that each case of exclusion fits the description of either those who are potential persons (in the sense of being potentially reasonable and rational) or those who will never become persons. The

way in which each case of exclusion is categorised is by focusing on the characteristics of the majority within each excluded group, as well as on the boundaries of current biological and medical technologies.

More specifically, the first part of my thesis, which addresses the question of exclusion as it applies to those who have the potential to develop the two moral powers, focuses on children, foetuses, and unreasonable persons. In the second part of the thesis I examine the implications of Rawlsian contractualism for those who will never develop the two moral powers by focusing on the exclusion of individuals with severe cognitive disabilities and of non-human animals.

1.7.3 The Problems Generated by Exclusion

I have argued that the intuitive argument about exclusion in social contract theories applies to Rawlsian contractualism. Given that those in the justificatory constituency are reasonable and rational, the exercise of political power need not be justifiable to those who are *not* reasonable and rational in order to be legitimate. Similarly, given that those in the Original Position are rational, mutually disinterested individuals concerned to advance their own conceptions of the good, as the theory stands, we may presume that their decisions will not take into consideration the interests of those excluded. Lastly, if principles of justice are to determine the terms of cooperation, we may doubt whether the principles apply to non-cooperators; indeed, this seems to be why the principles do not apply to non-human animals, for example.

I contend that the exclusion of these groups from the constituencies of justice and legitimacy generates a number of pressing problems for Rawlsian contractualism. Some of these

problems, such as the difficulties that the theory faces in giving an answer to the question of abortion, have already been identified by critics, albeit not as problems generated by the Rawlsian account of political personhood and social cooperation in particular. Thus, apart from aiming to resolve the question of exclusion in Rawlsian contractualism, this thesis further aims to reframe some of the problems that the theory faces as problems of exclusion.

The first problem that the question of exclusion gives rise to is that it threatens both political liberalism and justice as fairness with internal inconsistency. To see why, consider Rawls's solution to the problem of intergenerational justice. As we have seen, although the parties know that they represent individuals who are full persons in the present, they are required to imagine that all previous generations would have followed the principles that they choose. This solution is given in order to avoid what I call the *problem of temporal partiality*: any bias in favour of individuals who happen to exist at the point that the (real or hypothetical) social contract takes place, merely because they happen to exist at that point in time, would be unfair. Yet the question of exclusion highlights that Rawlsian contractualism itself is vulnerable to the charge of temporal partiality. For, even if the question of future generations' exclusion has already been resolved in the way Rawls recommends, there are a number of groups with respect to whom this question remains – namely, foetuses, children, and unreasonable persons who have the potential to become reasonable. So, even if future generations constitute an exception, the conditions for legitimacy and justice seem in other ways biased in favour of those who happen to count as persons at the point that the hypothetical contract or justification takes place. The first problem that the question of exclusion poses to Rawlsian contractualism is, therefore, the charge of internal inconsistency: for this question reveals instances of temporal partiality to which Rawlsian contractualism is theoretically opposed.

A second problem generated by exclusion is the threat of *indeterminacy*: that is, the theory may be unable to defend a specific position on certain political issues. Although this problem has already been identified in the academic literature, its proponents have not clarified that the primary source of the problem is the question of exclusion. Framed in this way, the problem of indeterminacy highlights that, because certain possible individuals, such as foetuses and future generations, are excluded from the constituencies of justice and legitimacy, the implications of Rawlsian contractualism for the policies that concern them are problematically indeterminate. And due to the commitments of public reason, Rawlsian contractualism cannot appeal to comprehensive doctrines, in the form of, say, a metaphysical conception of personhood, to defend a specific view or policy on these issues. In fact, critics suggest, Rawlsians are unable to take a stance on these issues without implicitly relying on a comprehensive doctrine. As a result, it is often argued that political liberalism is indeterminate, and hence objectionably incompatible with *all* positions on certain important issues.

Finally, the third problem generated by the question of exclusion is that, in some cases, Rawlsian contractualism gives rise to counterintuitive implications. Even if one thinks that excluding non-human animals from the constituencies of justice and legitimacy is not problematic, the fact that some *humans* are not counted as free and equal, are excluded from these constituencies, and may not be granted any protection of their interests in the basic structure of society does seem objectionable. Apart from the external objection that these implications are deeply problematic in themselves, Rawlsian contractualism further faces another internal objection of inconsistency. This is the case because, according to Rawls, the implications of his theory for public issues must 'lie within the leeway allowed by each

of the reasonable comprehensive doctrines making up an overlapping consensus' (Rawls 2005, 246). But if the theory has counterintuitive implications, it clearly fails to satisfy this requirement. We therefore see that the third problem that arises due to the question of exclusion is that of counterintuitive implications; and this problem is important for reasons both external and internal to the theory.

This is only a brief overview of some of the problems that are generated by exclusion in political liberalism. I discuss them and others in more detail in subsequent chapters, examining how they apply to each particular case of exclusion. Before doing so, however, I want to argue now that they cannot be solved by any general approach but must be addressed on a case-by-case basis.

1.7.4 Possible Solutions

It might be thought that all these problems can be addressed all at once by simply revising one or another particular feature of Rawlsian contractualism. If this were the case, then there would be no need to consider each issue separately; rather, the areas of tension that I have identified could be eliminated much more simply. Recall that, according to Rawlsian contractualism, the content of justice and political legitimacy is given by imagining the terms of the social contract that would be agreed to by idealised individuals under idealised conditions. The obvious suggestion, then, would be that we change either the conditions under which the contract is made or the way in which the parties are idealised.

Indeed, the former seems to be Rawls's methodology in his solution for the issue of future generations. As we have seen, in order to address the problem of temporal partiality in the context of intergenerational justice, he argues, we must require the parties to agree to

principles knowing that the same principles would have applied to the generations preceding them as well. It might then be suggested that the same solution could be used to respond to all cases of exclusion.

However, even if this refinement resolves the problem of temporal partiality and the related problems that it generates for policy issues that concern future generations, it cannot be used for other problems of exclusion, such as the cases of individuals with severe cognitive disabilities and non-human animals. This is because these groups' exclusion is not merely an issue of time: unlike members of future generations, who *will* develop the two moral powers,[11] these groups will *never* become persons in the political sense. In fact, the solution to the problem of future generations is not even sufficient to resolve the problem of temporal partiality as it applies to cases other than strictly speaking future generations, such as foetuses and children. To see why, imagine that the parties were asked to decide on the state's policy regarding contraception on the condition that the generations before them would have followed the same policies. In this scenario, it seems likely that the parties would have chosen to ban contraception out of fear that they might end up not being brought into existence.[12] Of course, there is more to be said about this possible line of argument, which I discuss later on in the thesis. The general point, however, is that the solution that Rawls offers for the problem of intergenerational justice does not seem suitable to resolve all other problems of exclusion.

Another change in the conditions of the hypothetical contract that, it might be thought, would resolve the problems of exclusion would be to abandon the commitment to social

[11] Of course, given the non-identity problem, this claim is presented in a simplified form. Justi *who* will exist and develop the moral powers is a matter of precidely what those who currently count as persons do.
[12] For a similar argument regarding foetuses, see Evers 1978, 111; Shaw, 2011, 96.

cooperation or the role of mutual advantage as the primary aim of social cooperation. It might be argued that this would resolve all problems of exclusion because it would imply that groups which cannot participate in or make significant contributions to social cooperation could be fully included in the theory. However, this does not seem to be a viable refinement either, as it would require changing the very core of Rawlsian contractualism. The resulting theory of justice and legitimacy would inevitably be a rejection, not a refinement of Rawlsian contractualism.

A different possible solution to the problems of exclusion would be to change the ways in which the parties to the hypothetical contract are idealised. One way in which the idealisation of the parties could change would be to assume that the parties in the Original Position are not just rational, but reasonable as well.[13] The suggestion here is that, if the parties were not solely motivated by their own advantage, they would come to consider the interests of all those who are excluded from the constituencies of justice and legitimacy. However, this revision would seem question-begging, as it would require us to assume that the parties are motivated by their sense of justice; yet, given that the Original Position is a mechanism designed to uncover what justice requires, assuming that the parties are reasonable would require us to presuppose too much about the content of justice.

It might be suggested that the problems of exclusion would be resolved if we changed the criteria for political personhood. By dropping the role of reasonableness and rationality in the theory and by endorsing a more inclusive conception of personhood, a greater number of groups would be included in the constituencies of justice and legitimacy. Recall,

[13] Recall that the Original Position and the Veil of Ignorance represent the reasonable, while the parties are only viewed as rational, mutually disinterested persons.

however, that reasonableness and rationality are chosen as the characteristics of political personhood because only individuals who possess them are capable of participating in social cooperation conceived as having the main purpose of mutual advantage. As a result, it seems unlikely that a different account of personhood could preserve the distinctive role of social cooperation and mutual advantage in Rawlsian contractualism; and as we have seen, an account that would not preserve that role would require us, not simply to refine Rawls's theory of justice and legitimacy, but to dismiss it.

Finally, it might be argued that the term *capacity* should be interpreted in terms of *potentiality*. This claim was in fact endorsed by Rawls: the thought is that moral personality is 'potentiality that is ordinarily realized in due course' (Rawls 1971, 442). The objective behind this claim seems to be the inclusion of children in the realms of justice and legitimacy. However, despite Rawls's own endorsement of this claim, it is not possible to resolve the problems of exclusion in this way. To begin with, even if we accepted it as a solution, some cases of exclusion would remain unresolved. For this argument would not ensure the inclusion of those who lack the potential to develop the two moral powers – namely, individuals with cognitive disabilities and non-human animals. Moreover, I suggest that this solution is not plausible, even if we only focus on the cases in which it does apply. This is because this solution would imply that, not only children, but also foetuses and future generations should count as *full* free and equal persons with rights and interests that matter as much as those of those who are currently reasonable and rational. Therefore, the claim that the capacity to be reasonable and rational should be understood in terms of one's *potential* to become reasonable and rational is both unable to resolve all problems of exclusion and independently implausible.

We can thus conclude that the problems of exclusion cannot be *jointly* addressed by one major revision of the conditions of the hypothetical contract or of the political conception of the person. It follows, then, that each of these problems needs to be explored separately. This is what I do in the remainder of the thesis.

1.8 Thesis Outline

More specifically, the first part of the thesis focuses on the exclusion of those who have the *potential* to develop the two moral powers if current persons act (or refrain from acting) in certain ways. In the second chapter, I begin by examining how the problem of exclusion applies to children. I argue that the definition of political personhood implies that children are excluded from Rawlsian contractualism in three ways. First, they are excluded from the constituency to whom justification for the exercise of political power is owed. Secondly, their interests are not taken into consideration in the Original Position. And thirdly, it is doubtful whether the principles of justice that are collectively chosen in the Original Position, such as principles that protect basic rights and liberties, apply to children. In the second section, I propose a way of answering the question of exclusion in the case of children, without abandoning the theory's core conceptions of personhood and social cooperation. I argue that there are strong reasons to hold that all humans, including children, have an interest in developing the two moral powers. I then show how this proposal can resolve the problem of children's exclusion by incorporating it into the Original Position.

In Chapter 3, I turn to the question of foetuses' exclusion. I argue that, contrary to what is typically held, the implications of Rawlsian contractualism for the question of abortion do not undermine the theory in any way. In doing so, I respond to three distinct arguments, which purport to show the opposite. According to the first argument, Rawlsian

contractualism is incomplete: it is either indeterminate or inconclusive on the issue of abortion because it does not permit appealing to comprehensive doctrines that include, for instance, metaphysical conceptions of personhood in order to provide us with an answer on the permissibility of abortion. The second argument states that, precisely because public reason liberals cannot appeal to non-political conceptions of personhood, the only consideration available to determine the political liberal view of abortion is women's rights and interests. Thus, the argument goes, Rawlsian contractualism has radically permissive implications for the legal regulation of feticide, even in late-stage pregnancies. On the contrary, according to the third argument, public reason implies that abortion is *never* permissible because those deciding on the principles of justice in the Original Position would want to safeguard themselves from ending up as aborted foetuses. By showing how each of these arguments fails, I defend Rawlsian contractualism against the charge that its implications for abortion undermine it.

In the fourth chapter, I argue that, as Rawlsian contractualism stands, it seems to exclude unreasonable persons from the constituency to whom justification for the exercise of political power is owed and from the constituency to whom the principles of justice apply. I suggest that unreasonable persons *can* be included in the theory once we notice three important reasons why states have a duty to enable them to develop their reasonableness. These reasons are that they have a basic interest in developing their two moral powers, that states ought to cultivate their citizens' reasonableness in order to protect their other citizens' basic rights and interests, and that states ought to cultivate reasonableness in order to promote stability for the right reasons. I then examine the policy implications of my argument and claim that certain controversial policies such as mandatory education,

therapy, and neurointerventions are sometimes required for the state to fulfil its duty to cultivate the reasonableness of unreasonable persons.

The second part of the thesis focuses on the exclusion of those who will *never* sufficiently develop the two moral powers. The first chapter of this part (Chapter 5) explores the problem of exclusion as it applies to individuals with severe cognitive disabilities. The purpose of this chapter is to respond to show why and how Rawlsian contractualism is not exclusionary with regards to the rights and interests of individuals with cognitive disabilities. In doing so, I begin by explaining why, contrary to what has been suggested by some critics, the problem of exclusion applies *only* to individuals with severe *cognitive* disabilities and not to individuals with physical disabilities. I then examine a number of solutions that have been offered to this problem by a number of philosophers and show why they fail to produce a more inclusive account of Rawlsian contractualism. In the final section of the chapter, I defend a twofold solution. First, I claim that many individuals with cognitive disabilities may have the potential to develop the two moral powers and therefore an interest in satisfying this potential, which ought to be taken into consideration by the members of the constituencies of justice and legitimacy. Whenever it is not clear whether one does have this potential and the relevant interest, states and citizens should err on the side of caution and treat these individuals as they would if they did have the relevant potential and interest. Second, I argue that creating children gives rise to special parental duties and related societal duties towards children and that these duties apply to children with severe cognitive disabilities. I suggest that this is because those included in the justificatory constituency ought to show opacity respect to all children and parents and that the only way to do so is to refrain from assessing a child's moral powers when it comes to determining whose interests ought to be fulfilled.

Finally, in the last chapter of the thesis, I explore how the question of exclusion applies to non-human animals. In the first section of the chapter, I examine and reject solutions that have already been offered in response to the question of exclusion. Briefly, these solutions are an intuitive equality argument; the claim that we have relationships with animals for which the two moral powers are not required; the argument that contractualism could include an appeal to altruism; an argument from overlapping consensus; an argument from capabilities and species membership; and an intuitive argument in favour of trusteeship. After rejecting these potential solutions to the problem of exclusion, I offer an alternative solution. More specifically, in the second part of the chapter, I argue that there are duties of compassion that establish a *pro tanto* reason to take animals' basic interests into consideration in the justificatory constituency's deliberation. These duties exist because, given our understanding of human psychology, the cultivation of citizens' reasonableness requires the cultivation of a concern for all sentient beings' basic interests. Although these duties of compassion are less weighty than the duties that I have defended towards other groups that may seem to be excluded from Rawlsian contractualism, they nonetheless address the question of exclusion in a way that prevents the theory from having extremely counterintuitive implications.

Chapter 2

Children

2.1 Introduction

As I explained in the first chapter of the thesis, there are two different categories of individuals to whom the question of exclusion applies in distinct ways. The first category includes those who have the potential to develop the two moral powers and, by extension, to participate in society conceived as social cooperation for mutual advantage, as soon as their moral powers reach the necessary threshold. This potential can be realised if and only if those who fulfil the political personhood requirements at the present moment act (or refrain from acting) in specified ways. The second category is comprised of those who will never develop the two moral powers to the degree that is associated with the status of political personhood.

In this chapter, I explore the problem of exclusion as it applies to one group within the first category, namely children. In particular, I argue that the definition of political personhood, combined with the significance of social cooperation in Rawlsian contractualism, implies that children are excluded from the constituencies of legitimacy and justice.[14] This argument is developed in the following section, where I show that the question of exclusion applies to children in three ways: first, they are excluded from the constituency to whom justification for the exercise of political power is owed; second, their interests are not taken

[14] This issue has not been addressed by political liberals, who often assume that the political liberal state has certain duties towards children and proceed to show the implications of these duties (see Fowler 2010; Clayton 2006).

into consideration in the Original Position; and lastly, there are reasons to doubt whether the principles of justice apply to children.

As I argued in the first chapter, if the question of exclusion remains unanswered, Rawlsian contractualism faces three pressing problems. The first problem is that the exclusion of certain groups, including that of children, makes Rawlsian contractualism both internally inconsistent and independently objectionable. To see why, recall that the exclusion of certain parties reveals temporal partiality. This is the case because the way in which political power is justified as legitimate is unfairly biased in favour of those who happen to count as persons at the point that the (real or hypothetical) contract takes place. Most children are clearly among those who are disadvantaged in this sense, given that they have not yet developed the two moral powers to the appropriate degree. The fact that their exclusion is the outcome of temporal partiality brings us to the problem of internal consistency: as I argue in the first chapter, Rawlsian contractualism is theoretically opposed to temporal partiality as is shown by Rawls's recognition of the need to avoid it in connection with the treatment of future generations. Rawls states that the parties in the Original Position would 'be required to agree to a savings principle subject to the further condition that they must want all *previous* generations to have followed it' (Rawls 2005, 274). Therefore, unless it can be demonstrated that the exclusion of children does not generate temporal partiality, Rawlsian contractualism is threatened by internal inconsistency. For it may turn out that it incorporates temporal partiality while also accepting that this kind of partiality is unfair. And to the extent that temporal partiality is indeed an instance of unfair bias, Rawlsian contractualism is further vulnerable to the external criticism that it is an implausible theory of justice and legitimacy.

The second problem that arises from the question of exclusion, as we have seen, is the threat of indeterminacy. In the case of children, the worry is that, *because* they are excluded, Rawlsian contractualism cannot defend a specific position on the question of what duties, if any, are owed to children.[15] This is due to the theory's reliance on political liberalism and the related idea of public reason, which require that issues of basic justice and constitutional essentials be justified in political, rather than comprehensive, terms. Yet, the argument goes, if children are not persons in a political sense, and if political liberalism can reach conclusions only for political persons, the political liberal state cannot reach any conclusions on matters that concern those who are not political persons. Doing so would necessarily involve appealing to a comprehensive doctrine, such as a metaphysical conception of personhood, on the validity of which there is reasonable disagreement. Thus, the question of children's exclusion leaves Rawlsian contractualism vulnerable to the charge of indeterminacy.

Finally, precisely because children may be subject to certain objectionable kinds of exclusion, such as the ones that I explicate in the next section, Rawlsian contractualism faces the objection that it has extremely counterintuitive and undesirable implications. For example, the exclusion of children might imply that they do not possess any rights and that neither their parents nor society have any duties towards them. Clearly, this implication would be viewed as undesirable and unacceptable from the viewpoint of all reasonable comprehensive doctrines. Given that Rawls states that the implications of his theory must 'lie within the leeway' of citizens' reasonable comprehensive doctrines, it seems that he would acknowledge the force of this objection (Rawls 2005, 246). Thus, if children are

[15] Although it has not been argued, to my knowledge, that the exclusion of children may give rise to indeterminacy, a number of philosophers have discussed the general issue of indeterminacy in political liberalism (see e.g., Sandel 2005; Shaw 2011; Friberg-Fernros 2010).

excluded in the ways that I describe in the next section, Rawlsian contractualism is vulnerable to all of these three objections.

For that reason, there is a pressing need for Rawlsians to explore how their theory can include children or accommodate their interests. This is what I do in the third section. In particular, I propose a way of responding to the question of exclusion without abandoning the ideas of political personhood and social cooperation, which are central in Rawlsian contractualism. I argue, consistently with certain of Rawls's remarks in *Political Liberalism* and *Justice as Fairness*, that there are strong reasons to hold that all humans have an interest in developing the two moral powers. I suggest that, by focusing on this interest and by incorporating it into Rawlsian contractualism through a refinement of the Original Position, the theory can recognise certain important interests of and duties to children, thereby responding to the question of exclusion.

The final part of this chapter explores the implications of my argument for a political liberal state's child-centred policies. In particular, these policies concern the ways in which parental and citizens' duties towards children ought to feature in the legal structures and public policies that concern education and the family. It should be noted that these implications are taken to be valid for all theories within the family of liberal conceptions of justice that are consistent with the requirements of political liberalism, and are therefore not limited to justice as fairness.

2.2 The Exclusion of Children

Recall that, due to its commitment to political liberalism and public reason, Rawlsian contractualism defines personhood in political rather than metaphysical terms. The

resulting account of personhood defines persons as those who possess the moral powers of reasonableness and rationality to the appropriate degree. As I argued in the first chapter of the thesis, the reason why these two properties are viewed as political rather than comprehensive is that they are necessary for one's meaningful participation in social cooperation with the purpose of mutual advantage. On the one hand, reasonableness ensures that one is willing and able to cooperate on fair terms. On the other hand, being rational implies that one has interests that can be advanced through cooperation with others in a system designed for *mutual* advantage.

Although the idea of personhood is distinct from the idea of citizenship, it is already clear that the two are connected, because the former is a necessary condition for the latter. That is, in order to be viewed as a citizen, one must meet the necessary (even if not sufficient) condition of counting as a person in the political sense.[16] For instance, non-human animals do not count as citizens, even if they live within the relevant territory, *because* they do not possess the capacities that are associated with political personhood. This is because the idea of citizenship is connected to the idea of social cooperation in the way that the idea of political personhood is: being a citizen means that one is capable of participating in social cooperation. Although a full analysis of all the conditions for citizenship would be both interesting and important, it is unnecessary for my purposes. The relevant point is that those who do not fulfil the personhood criteria cannot fulfil a central citizenship criterion. This is important because, as we will see, some conditions for inclusion in Rawlsian contractualism as it stands refer to the idea of a free and equal *citizen*.

[16] This condition might not be sufficient because there may be further conditions that have to be met, such as being resident within the relevant territory.

It seems quite straightforward then, that, in Rawlsian contractualism, children do not count as persons in the political sense, nor can they be viewed as citizens.[17] As I will now argue, this implies that children are excluded from the political community in three ways. First, they are excluded from the constituency of legitimacy – that is, the constituency to whom political power ought to be justifiable. Second, their interests are not accommodated by Justice as Fairness. Third, they seem to be excluded from those to whom the principles of justice apply.[18]

2.2.1 Exclusion from the Justificatory Constituency

To see why children are excluded from the justificatory constituency, recall the liberal principle of legitimacy:

> our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all *citizens as free and equal* may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason (Rawls 2005, 137, my emphasis).

Given that the exercise of political power ought to be justifiable to citizens, and given that children do not count as citizens, they are excluded from the justificatory constituency. This means that the exercise of political power may be legitimate, even if it is not justifiable from children's point of view.

The exclusion of children from the justificatory constituency may seem objectionable for both intrinsic and extrinsic reasons. It may be objectionable in itself if being included in

[17] I address one obvious objection to this – namely, that Rawls thinks of his principles as applying to people over their whole lives, not just from the age of reason – below.
[18] It should be stressed that, although these are three different *ways* in which children are excluded, my argument is not that there are three different *sources* of exclusion. In all cases, the source of exclusion is Rawls's conception of political personhood (and, by extension, his conception of citizenship).

the justificatory constituency is *intrinsically* valuable. That is, we may think that, even if one's interests are represented and protected in a well-ordered society, one's exclusion from the (hypothetical or real) constituency to whom political power is justified is objectionable. This is because we typically view *participation* – and not merely *representation* – in political matters as an issue of symbolic importance that signifies our relations as *equals*. However, the force of this objection is discounted by the fact that children will normally be included in the justificatory constituency in the future, once they develop the two moral powers to the sufficient degree. Thus, even if we accept that there is something *pro tanto* wrong in the exclusion of children from the justificatory constituency, this exclusion may be all-things-considered acceptable.

However, there is a second and more challenging problem that may arise from this type of exclusion: political power may be viewed as legitimate even if it is exercised in a way that fails to protect or harms children. This is because being included in the justificatory constituency is *extrinsically* valuable, as it ensures that one's interests are represented and therefore taken into consideration in the exercise of political power. Those excluded, such as children, face the risk of their interests being discounted or excluded altogether. Of course, someone might object that the 'principles and ideals acceptable to their [i.e. citizens'] common human reason', which the liberal principle of legitimacy requires citizens to consider in the process of justification, would prevent them from endorsing such implications. Indeed, it seems uncontroversial that there is overlapping consensus among all reasonable comprehensive doctrines on the fact that children's development and well-being ought to be promoted. However, it is still the case that any policies or laws that concern children must be justifiable only to citizens: even if they are justifiable to children and the adults the children will become as it turns out, they need not be. And in practice,

given that citizens' interests, either qua citizens or qua parents, are sometimes inconsistent with children's interests, we may conclude that certain decisions reached will be different from the ones that would be reached if children were included in the constituency to whom justification is owed.[19]

2.2.2 Exclusion from the Represented Constituency in Justice as Fairness

The second sense in which children are excluded in Rawlsian contractualism, which follows from the first, concerns the fact that the principles of Justice as Fairness are designed in a way that does not take their interests into account. This is due to the design of the Original Position, which states that only reasonable and rational persons are to be represented and that they are to be portrayed as mutually disinterested, rational individuals. These two methodological constraints imply that, firstly, children are not represented in the Original Position and that, secondly, those who are represented are not interested in how the distribution of the benefits and burdens of social cooperation affects others, including children. In the words of Williamson Evers, 'because parties to the agreement in the Original Position are adults and must know that they have adult rationality, it cannot correctly be said of them that "no one knows his situation in society, … and therefore no one is in a position to tailor principles to his advantage"' (Evers 1978, 110, quoting Rawls 1971, 139). Of course, this type of exclusion is related to exclusion from the justificatory constituency, given that those represented in the Original Position are the same parties who are included in the justificatory constituency. Yet, although the two types of exclusion are closely related, they are distinct in that the first refers to legitimacy and the second refers to justice.

[19] One such example is that parents may choose to induct their children into a faith via a painful procedure (Fowler 2010, 374). Another example is the climate and ecological breakdown that our societies are currently heading for, which the school strikes movement is currently resisting.

Due to the exclusion of children from the constituency that is represented in Justice as Fairness, it becomes possible that the benefits and burdens of social cooperation are distributed in a way that burdens children more compared to a different distribution in which children would have been represented on a par with reasonable and rational persons. When judged from the perspective of children's interests, the principles of justice would therefore be inferior to an alternative set of principles that would take those interests into account. Moreover, it seems possible that the principles of justice chosen could *harm* children. This is because the content of the principles might not only be *inferior* to other principles, from the perspective of children's interests, but actually *opposed* to children's interests, as for example in cases where children's interests clash with the interests of those fully included.

Someone might respond that, by protecting individuals' basic liberties and equal opportunities and by implementing the difference principle, the procedure of the Original Position protects children's basic rights and interests in exactly the same way that it protects citizens' basic rights and interests. However, this seems to presuppose too much, as it assumes that the principles of justice apply to non-persons. Yet if the principles do not apply to non-human animals and foetuses, precisely because non-human animals and foetuses are not sufficiently reasonable and rational, it is not clear why they should apply to children, who similarly lack the two moral powers.

2.2.3 Exclusion from the Constituency to Which Justice as Fairness Applies

This brings us to the third way in which children are excluded from Rawlsian contractualism as the theory stands: there are good reasons to doubt whether the principles

of justice can even apply to children, given that the principles apply only to citizens and it is doubtful that Rawlsians could view children as citizens. Although this type of exclusion may seem similar to the previous one – namely, that children's interests are not represented in the constituency of justice – there is an important difference: the claim here is not that the principles chosen do not protect children's interests but, rather, that even the principles that could protect children's interests if they applied to them do not, in fact, do so because they do not apply to children in the first place. If left unchallenged, this type of exclusion would imply that children are not even to be granted the protection of the rights and liberties that the first principle of justice guarantees; nor are their opportunities or access to resources to be targeted by the state in the way that the second principle of justice requires.

At this point, Rawlsians might object to the assumption that the possession of the two moral powers is a *necessary* condition for citizenship; instead, they might claim that children are citizens simply because they have been born in the well-ordered society of justice as fairness. Indeed, Rawls states that because society is assumed to be closed,

> entry into it is only by birth and exit from it is only by death. We are not seen as joining society at the age of reason, as we might join an association, but as being born into society where we will lead a complete life (Rawls 2005, 41).

Thus, it might be argued that the exercise of political power ought to be justifiable to children and that the principles of justice should apply to them because the only criterion for counting as a citizen is to have been born into the well-ordered society and to have the *prospect* of leading a complete life in it.

However, even if children are to be viewed as citizens in virtue of their birth, the second kind of exclusion may still obtain: that is, their interests may still go, at least partially, unrepresented in the theory of justice that guides the distribution of benefits and burdens in the well-ordered society. This is because the parties in the Original Position represent sufficiently reasonable and rational persons – and children do not typically possess these capacities. As a result, the parties will effectively assume that those that they represent have already reached adulthood. Therefore, even if those the parties represent count as citizens over a complete life, the adult-part of their life will be represented more strongly than the part of their childhood. The interests of children qua children would thus be excluded.

A different way of making this point is that, even if children count as citizens, they cannot be viewed as *equal* citizens, given that the basis of equality is specified as the possession of reasonableness and rationality to the appropriate degree. Thus, because children do not meet the equality test and because the liberal principle of legitimacy requires that the exercise of political power be justifiable to 'all citizens as free *and equal*', we can conclude that, even if children do count as citizens, they are not included in the justificatory constituency insofar as they are children. In order to claim that *all* citizens are free *and equal* and that everyone who has been born into society is a citizen, Rawlsians would need to abandon the role of reasonableness and rationality as specifying the basis of equality.

Yet making this move would raise the following problem. In order to count as an equal citizen, one would not need to possess the two moral powers which are necessary in order to participate in social cooperation; one would merely need to have been born into society. The basis of equality would thus be birth. However, the mere fact that someone has been born seems morally arbitrary as a determinant of their moral and political status: there don't

seem to be any morally relevant differences between an infant right before it is born and an infant that has just been born. Similarly, plenty of non-humans are born within the territories of the relevant society, yet this does not affect their moral and political status. Rawls's own writings are not helpful here because he states that one counts as a citizen from the moment of birth without offering a justification for this claim. It seems that the only justification that he could have given is that birth ensures that one will develop the capacity to participate in social cooperation.

More precisely, because humans lack the relevant capacities when they are born, it seems that what matters, if we accept the birth condition, is that they have the *potential* to develop the two moral powers and participate in social cooperation. But if having this potential is all that matters, then it is not clear why birth should be viewed as the relevant proxy for this potential. For example, foetuses, zygotes, and future generations seem to have exactly the same potential. As a result, if Rawlsians were to change the basis of equality in the suggested way, it seems that they would have to accept that all beings who merely have the potential to develop the powers that are necessary for one's participation in social cooperation are equal citizens. Yet the claim that those who merely have the potential to develop the two moral powers, such as zygotes, are *equal citizens* seems both counterintuitive and problematic in terms of its implications for a liberal theory that seeks to defend laws and policies in areas such as access to abortion and stem-cell research. Consequently, Rawlsians cannot resolve the problem of exclusion by stating that children count as citizens just because they have been born.

Thus we return to my claim that Rawlsian contractualism seems to exclude children in three ways. First, they seem to be excluded from the constituency to whom justification for

the exercise of political power is owed, as that is specified by the liberal principle of legitimacy. Second, they seem to be excluded from the Original Position, which implies that the principles may be designed in a way that does not take their interests into consideration. Third, they seem to be excluded from the constituency of justice in the sense that the principles of justice do not apply to them. As I have argued, if left unchallenged, the exclusion of children generates the pressing problems of internal inconsistency, indeterminacy, and undesirable consequences for Rawlsian contractualism.

2.3 Avoiding Exclusion

An unsatisfying solution would be to endorse Rawls's *first* solution to the problem of future generations. In *A Theory of Justice*, Rawls suggests that the parties in the Original Position represent heads of families, as opposed to individuals. This is presented as a solution to the problem of intergenerational justice because it is stipulated that all heads of families have 'ties of sentiment between successive generations' (Rawls 1971, 292). Thus, Rawls argues that the parties would agree to a just-savings principle because they would care about securing a fair share of advantages for their successors.

There are two reasons why someone might suggest that this solution for the exclusion of future generations would resolve the problem of children's exclusion too. First, the members of future generations who are closest to any given person might naturally be supposed to be that person's own children (Brennan and Noggle 2000, 50). Second, as outlined in the introduction, the exclusion of children is similar to the exclusion of future generations in that they both belong to the group that is excluded because they have the *potential* to develop the two moral powers if those who currently count as persons act or refrain from acting in certain ways. Since neither children nor future generations count as

persons in the present, and since political liberalism cannot appeal to a metaphysical conception of personhood that distinguishes future generations from foetuses, and foetuses from children, these groups have the same status of non-personhood. Thus, we may expect a solution to the problem of future generations' exclusion to similarly resolve the problem of children's exclusion.

However, as many critics have already pointed out, the solution offered by Rawls in *A Theory of Justice* to the issue of intergenerational justice is deeply unsatisfying. To begin with, the solution rests on the controversial assumption that all heads of families would have emotional ties with future generations, many of which they would never come to meet. Given that different individuals experience extremely different ties of sentiment to future (and past) generations, the grounds on which Rawls bases his solution appear shaky (Brennan and Noggle 2000, 48). Perhaps an even more important critique is the one that has been put forward by feminists, such as Susan Moller Okin: if the parties in the Original Position represent heads of families as opposed to individuals, then this implies that the resulting theory of justice would not address issues of justice *within* families, which, in our historical context, disproportionately disadvantage women and children (Okin 1989, 95). For instance, given that the parties would know that they represent *adult* heads of families, they might not take children's interests into consideration. We may therefore conclude that Rawls's first proposed solution to the problem of future generations (and by extension to the problem of children) fails to resolve the problem of exclusion.

2.3.1 The Interest in Developing the Two Moral Powers

I contend, however, that Rawlsian contractualism can avoid the exclusion of children and its related problems by stressing that all humans have an interest in developing the two

moral powers.[20] In this section, I show how this argument is in fact implied by Rawls's theory and argue that it ought to be factored into the Original Position. Once the Original Position is modified in the appropriate ways, the theory gives rise to plausible implications for child-centred law- and policy-making.

I suggest that all humans have an interest in developing the two moral powers because they receive two distinct benefits from their participation in social cooperation. Given that these benefits are *political* (since they are justified by reference to social cooperation), the interest in developing the two moral powers, which is justified in virtue of these benefits, is a *political* interest. First, participating in social cooperation is *instrumentally* good for persons, as it is the best means for each person to secure the end of their individual advantage, which is specified by each person's conception of the good. This further implies that participating in social cooperation is also the best means to secure all persons' mutual advantage.[21] Since individuals have an interest in participating in social cooperation with the purpose of mutual advantage, and since doing so requires possessing the two moral powers to the requisite degree, it follows that individuals have an interest in developing the two moral powers.

Second, and more controversially, citizens' participation in social cooperation is *intrinsically* valuable, because the exercise of the two moral powers is intrinsically good. Exercising the two moral powers by participating in social cooperation ought to be among

[20] It might be objected that, if this is true, then zygotes and foetuses would have the same interest and that this would lead to illiberal conclusions. I consider this objection in the next chapter, which focuses on the exclusion of foetuses.

[21] Of course, this does not imply that any single individual views *mutual* advantage as one of her ends, although the fact that each cares about her own advantage does not imply that persons think in individualistic terms. For, citizens' conceptions of the good (and hence their conception of advantage) often include attachments to other individuals and groups.

individuals' *aims* and not merely a means to their other ends. This is because political liberalism is a *moral* conception, which ought to be affirmed as such, rather than as a modus vivendi. It thus requires that citizens view political society as intrinsically valuable and not simply as a contingently valuable arrangement from which they gain other advantages (Rawls 2005, xlii). If, therefore, exercising the two moral powers by participating in social cooperation ought to be among individuals' ends, and given that they have an interest in achieving their ends, we can conclude that they have an interest in exercising the two moral powers. Indeed, this seems to be implied by Rawls when he argues that social cooperation allows citizens to exercise their two moral powers, which is 'experienced as good' (Rawls 2005, 207).

Now someone might object that the first premise of this argument – the claim that exercising the two moral powers and participating in social cooperation ought to be among individuals' ends – is not consistent with Rawls's political liberalism and his account of public reason because it draws on comprehensive views on what is of value in life. However, one's exercise of the two moral powers is an intrinsic good in a *political* sense, precisely because it is defined by reference to social cooperation. That is, the social cooperation that occurs in the well-ordered society (as opposed to a *modus vivendi*) is possible only if citizens view their participation in it as intrinsically good. Hence, the claim that citizens ought to view society and their participation in it through the exercise of their moral powers as intrinsically valuable is not in tension with the commitments of political liberalism, but is in fact implied by them.

The claim that individuals have an interest in developing the two moral powers is not only implied by Rawls's (and, in fact, most contractualist theorists') arguments about social

cooperation, but it is also explicitly mentioned by him, although not fully analysed. For example, Rawls argues that the parties in the Original Position and the citizens they represent 'are motivated to secure their higher-order interests associated with their moral powers' (Rawls 2005, 74; cf. Scanlon 2004, 1481). Whether or not citizens develop the two moral powers is not viewed as neutral or as depending on luck; rather, the development of these powers is one of the aims of political justice. Similarly, Rawls recognises the implications of this interest for the way in which the motivation of the parties in the Original Position needs to be defined:

> the aim of the parties [in the original position] is to agree on principles of justice that enable the citizens they represent to become full persons, that is, adequately to develop and exercise fully their moral powers and to pursue the determinate conceptions of the good they come to form (Rawls 2005, 77).

Yet it might be objected that even if we accept that individuals have an interest in developing the two moral powers, this cannot be accommodated within the theoretical framework of the Original Position. This is because, the argument goes, if the parties in the Original Position are rational, mutually disinterested contemporaries who know that they represent persons who are *already* reasonable and rational, their rationality cannot motivate them to accept principles that ensure that future citizens *become* reasonable and rational. For their rationality and lack of interest in other persons' matters implies that they can only accept principles which further the interests of the people they represent. And if those that they represent are *already* reasonable and rational, they cannot possibly have an interest in *becoming* reasonable and rational, even if they had this interest before they developed the two moral powers. Thus, it may seem that, for the purposes of consistency, Rawlsians either have to dispute that we have an interest in developing the two moral powers, which would leave Rawlsian contractualism vulnerable to the problems of exclusion; or the Original

Position must be revised to ensure that the resulting theory of justice reflects our commitment to promoting a state's current *and* its future citizens' interests.

I contend that the latter can be achieved by following the *second* solution that Rawls has given to the problem of future generations' exclusion. That is, if we accept that non-persons have an interest in becoming persons in the political sense and that this interest ought to be promoted by the principles of justice, then this interest must be taken into consideration by the parties in the Original Position. Thus, we may place a constraint on the parties' choice of principles, according to which the resulting principles must be consistent with promoting future persons' interest in developing the two moral powers. In fact, this is already implied by Rawls's solution to the problem of intergenerational justice: if the parties can accept only principles that would have applied to previous generations and if they recognise that they once had an interest in developing the two moral powers, then they would only accept principles which would have ensured that they would become reasonable and rational, had they been followed by previous generations.

Now someone might object that in order to change the Original Position in this way, we must first make assumptions about which individuals with an interest in developing their moral powers are to be accommodated in the relevant way, and these will need greater defence than I have offered. For example, why shouldn't the interest that a foreigner may have in participating in a given scheme of social cooperation show that the Original Position has to be redesigned in cosmopolitan terms? Why should the interest that children have in participating in that scheme show that it should?

Before addressing this objection, it should be noted that those excluded can often provide those included with important benefits, which present those who are already included with good, mutually disinterested reasons to expand their constituency. For instance, many potential migrants can contribute to an otherwise closed society's production of primary goods. Similarly, society sometimes has an interest in parents' having and rearing children who end up developing the two moral powers. For example, society has an interest in its own orderly reproduction, given that a population decline would cause problems such as having sufficient taxpayers who could fund government spending for older generations.[22] We therefore see that those included in the justificatory constituency (and, by extension, the Original Position) would often have rational, mutually disinterested reasons to include *some* outsiders and to take their basic interests – such as the interest in developing the two moral powers – into consideration. However, in order to focus on the particularities of children's interest in developing the two moral powers, I set aside the question of whether there happen to be societal benefits in each case of potential inclusion.

Setting aside the potential societal benefits of a group's or individual's inclusion, the objection under discussion can be interpreted in different ways. If the worry is that the interests of those excluded do not suffice to show how those included ought to act, the response is that these interests may still present a *pro tanto* reason for acting in a way that takes those interests into consideration. For example, even if foreigners' interest in being included in the social cooperation of a specific state is not strong enough to warrant their inclusion (which is not a claim that I am defending here), their more fundamental interest in not being killed can impose at least a negative duty on those included to refrain from

[22] I return to this suggestion in Chapter 5. For more on the debate of children's contributions to society, see Casal and Williams 1995, 94; Folbre 1994, 86—90; Goodin 2005; Anderson 1999; Olsaretti 2013).

killing them. In general, the relevant interests only constitute one reason for action, which must be weighed against other reasons for action. Thus, my argument is not that the interest in developing the two moral powers is decisive; rather, I argue that there is a good reason in favour of satisfying it, which must be weighed against other reasons we may have.

The stronger form of this objection holds that those included do not have *any* duties towards those excluded, because the content of moral and political duties is determined by the decisions of those included. However, we cannot just assume that exclusion justifies the claim that we only have moral duties towards those included. In fact, my argument implies that this claim is unjustified, for if temporal partiality is unfair, then there is at least one broad group of excluded parties whose interests ought to be taken into consideration: namely, those who have the potential to become citizens with the passage of time.[23] Thus, the contractualist framework should not favour the citizens of the relevant society at Time A more than the citizens of the same society at Time B. Indeed, if the parties in the Original Position decided on principles permitting total depletion of the earth's resources, despite recognising that they had an interest in living in a world with natural resources, we would object that their decision is unfair. For, as Rawls's solution to the problem of intergenerational justice implies, if they, who are now persons, are glad that others respected their own interest before they were persons, they should be prepared to do the same for future generations. In this sense, given the theory's commitment to temporal impartiality, the idea of reciprocity must cut across times and bind different generations, instead of being confined to those who are currently citizens.

[23] Of course, it might be true that assuming that society is closed and self-sufficient is not only unrealistic, but also unfair. I am not making this claim, however, nor do I think that my argument against temporal partiality suggests that.

At this point, it might be objected that the appeal to our interest in developing the two moral powers will generate the implausible result that states should always try to increase and maximise their citizens' reasonableness and rationality via paternalistic policies. For example, states could promote the development of their citizens' moral powers by imposing mandatory education on adults who are already sufficiently reasonable and rational. This objection misfires, however, given that our interest in developing the two moral powers springs from our interest in participating in social cooperation; this latter interest implies that only a *threshold* of reasonableness and rationality is required, as no more than this is necessary for meaningful participation. This point highlights that the reasonable and rational capacities ground individuals' political status as equal persons solely because they are necessary conditions for social cooperation.[24] Indeed, this is why this conception of personhood is political rather than comprehensive, as the two moral powers are not defended as part of a perfectionist conception of what matters in life or on the basis of some deeper metaphysical truth. Thus, children's interest in becoming persons grounds an interest in developing the two moral powers *to a sufficient extent*. As a result, the duties that arise from this interest, such as the state's duty to offer education for all children, are only directed to those who have not yet reached the personhood threshold.

Nevertheless, some doubt may remain. For those pressing this objection may argue that if the two moral powers are indeed linked to individuals' capacity for social cooperation, we may suppose that some individuals will be able to offer a better quality of cooperation than

[24] My interpretation of Rawls's reasons for appealing to the two moral powers is very different from one that has previously been ascribed to him. According to Nathan, Rawls offers 'the intuited range property response': 'there is no more to say in justification of the importance of this level than "this is our practice," or "these are our values"' (Nathan 2011, 216). If my interpretation is correct, Rawls's appeal to moral personality is more like the 'substantive' view that Nathan ascribes, for example, to Hobbes, according to which there is an 'independent account of the significance of some threshold' (Nathan 2011, 216).

others (Brennan 2016). If this is the case, someone might argue that all individuals have an interest in participating in social cooperation *as well as possible*. This interest would then apply to all non-persons and persons alike, giving rise to objectionably paternalistic policies and discrimination on the basis of individuals' development of their moral powers.

However, I suggest that disregarding variations *above the threshold* in the degree of possession of the two moral powers is justified from within the contractualist framework. This is because those included in the justificatory constituency and, by extension, in the Original Position, would have good reason to decide that respecting individuals requires treating them as opaque by refusing to 'look inside them' (Carter 2013, 29). To see why, notice that those included have certain competing interests. On the one hand, they have an interest in developing their two moral powers, which suggests that variations *above* the threshold do matter and hence that they justify paternalistic interventions. On the other hand, the parties have an interest in leading an autonomous life: they have an interest in pursuing their own conception of the good without state interference, and they similarly have an interest in doing so *privately*, without facing judgment or assessment of their capacities by the state or other citizens. This interest in autonomy seems to pull towards a different direction than their interest in developing the two moral powers does. It thus seems plausible to conclude that the parties would settle the tension by agreeing on a *threshold* above which variation in individuals' moral powers would not matter from the state's point of view. Of course, this does not mean that variation above the threshold *never* matters. For example, certain relations, such as that between a therapist and a patient, require looking deep inside individuals in order to assess their mental state and how their well-being can be promoted. Similarly, relations based on merit, such as that between a teacher and a student, may require looking inside individuals to assess their relative

abilities, desires and motivation. However, when it comes to 'the relation between political institutions and citizens', viewing citizens as opaque if their moral capacities lie above the threshold seems appropriate (Carter 2011, 559).

So, my argument about our interest in developing the two moral powers is consistent with the claim that states ought to refrain from assessing a person's moral powers *above* the threshold that is necessary for social cooperation. Doing so would not be justifiable to reasonable and rational persons who want to protect and promote their interests, all things considered. This is even clearer when we take the Original Position to be the relevant contractualist method, for all parties would want to ensure that the state respects them by refusing to evaluate and target their two moral powers once they have developed them sufficiently. Therefore, we may conclude that focusing on the *threshold* of the moral powers is appropriate.

Now, once we accept that we have an interest in developing the two moral powers to the requisite degree, we are faced with a new difficulty: since each individual develops the two capacities at a different pace, it might seem that states ought to hold frequent examinations for all children, in order to determine whether their moral powers have reached the relevant threshold. For in the absence of such examinations, certain individuals who have developed their moral personality earlier than others in their age group have grounds to complain that they are disrespected and not treated as equal citizens (Fowler 2014, 101). Yet holding such examinations would be impractical, costly, and invasive. It would be impractical because it would be very difficult to come up with reliable and accurate tests for both reasonableness and rationality (Clayton and Chan 2006, 540; Fowler 2014, 103). This problem would be accentuated by the fact that we develop these capacities gradually, which makes it almost

impossible to pinpoint a specific point in time at which the threshold has been reached (Fowler 2014, 100). And even if that were possible, it would be extremely costly to test individuals every few hours, days, or months to determine whether they have reached the threshold. It seems, then, that there are separate considerations that pull to different directions. On the one hand, respect for individuals requires evaluating their cases on an individual basis when their moral powers are still below the appropriate threshold, in order to assess the kinds of duties that states have towards them. On the other hand, doing so raises questions of practicality, cost, and privacy.

This issue can be resolved by appealing to a constrained threshold that applies to age groups rather than individuals, albeit one that is endorsed as a *second-best* solution. This means that different age groups would be viewed as possessing the two moral powers to different degrees, depending on the average levels of development in each group (Fowler 2014, 95). Indeed, this seems to be the reasoning behind policies that require a minimum age, such as regulations about entering casinos. However, contra Fowler, who endorses this threshold on grounds of fairness, I suggest that whenever states know of exceptions without having to encounter issues of practicality, cost, and privacy, they should modify their attitude towards that person appropriately. For example, if there is undeniable proof that a child prodigy is sufficiently reasonable *and* rational, then the three factors no longer tilt us in favour of the constrained threshold and we may treat that person as an equal citizen. In these cases, paternalistic interventions such as mandatory education would be considerably more difficult to justify than they would be in standard cases regarding children. Similarly, there would be stronger reasons in favour of such interventions in cases of individuals who clearly and uncontroversially fall below the relevant threshold at an age where the majority lies above, when compared to cases of political persons.

2.3.2 The Interest in Developing the Two Moral Powers and Children's Inclusion

Recall that Rawlsian contractualism is vulnerable to the claim that children are excluded in three ways. First, it seems that they are excluded from the constituency to whom political power ought to be justifiable in accordance with the liberal principle of legitimacy. This might seem both intrinsically and extrinsically objectionable. The former worry is that being a member of the justificatory constituency is *intrinsically* good in the sense that there is a symbolic value in participating in that constituency as an equal and that, by being excluded, children are prevented from experiencing that good. The latter, more pressing worry is that those included in the justificatory constituency might view certain objectionable laws and policies as *legitimate*. Relatedly, this worry translates into a worry about justice: because children are not represented in the Original Position, their interests are not taken into consideration in the distribution of the benefits and burdens of social cooperation that the parties view as just. Third, there are reasons to worry that the principles of justice do not even apply to children. If left unchallenged, the exclusion of children gives rise to the problems of inconsistency, indeterminacy, and unacceptable implications. I contend, however, that recognising the interest that all children have in developing the two moral powers and ensuring that this interest is reflected in the design of the Original Position addresses the three charges of exclusion.

In particular, the claim that children miss out on an intrinsic good that is conferred by one's membership in the justificatory constituency loses its force once we notice that political power essentially ought to be justifiable to the future people that children will become. This is why those who are included in the constituency here and now ought to take future

people's (including children's) basic interest in developing the two moral powers into consideration. The charge that children's exclusion from the justificatory constituency is *extrinsically* bad is similarly addressed: for, even though they are not represented qua children, those who are included ought to promote children's interest in developing the two moral powers. Thus, no policies or laws that would harm children could be viewed as justifiable and, hence, as legitimate.

According to the second charge of exclusion, which follows from the first, the way that Rawls's own theory of justice, which is one theory in the family of liberal conceptions of justice, is built reveals a bias against children, leading to a distribution of the benefits and burdens of social cooperation that is likely to burden children more. This is because the principles are formulated by those who are represented in the Original Position, who, being mutually disinterested, only care about advancing their own interests. Given that children are not among those who are represented, it follows that their interests are not taken into consideration. However, even though children are indeed not represented in the Original Position, my argument in favour of considering individuals' interest in developing the two moral powers establishes that this kind of exclusion is not objectionable. For, given the relevant refinement of the Original Position, the parties do have to take into account children's interest in becoming persons in the political sense.

Finally, the third charge of exclusion – the charge of children's exclusion from the constituency to whom Justice as Fairness applies – states that the principles of justice do not apply to children and that, therefore, their interests cannot be promoted by appealing, say, to the difference principle or to the fair equality of opportunity principle. This argument is given by an analogy: if the principles do not apply to other non-persons, such

as non-human animals and foetuses, it follows that they cannot apply to children. Yet even if the principles of justice do not apply to children, the design of the principles must be consistent with children's developing the two moral powers. So, again, if this kind of exclusion exists, it is not necessarily objectionable. Thus, we see that the argument about children's interest in developing the two moral powers is sufficient to dispel our worries about the three kinds of exclusion.

2.4 Constructing Child-Centred Policies

Having examined how children can be included in Rawlsian contractualism, we face the question of what the practical impications of their inclusion are. In particular, Rawlsians must clarify what policies ought to be followed by political liberal states in order to ensure that children develop the two moral powers. Answering this question requires distinguishing between the relevant aims and effects of different policies. On the one hand, there are policies and laws that influence what we may call the *epistemic* aspect of reasonableness and rationality. These policies affect one's ability to understand the requirements of justice, as well as one's capacity to form and revise a conception of the good. On the other hand, other policies and laws affect the *motivational* aspect of reasonableness and rationality; that is, individuals' capacity to comply with the requirements of justice and to pursue their conception of the good. In this part of the chapter, I explore and suggest certain child-centred policies which target each of the two components of the two moral powers.

The third part of *A Theory of Justice* explores how children come to acquire a sense of justice in the well-ordered society. According to Rawls, who mostly follows Piaget, children go through three stages, which lead them to the development of reasonableness.

First, they adopt 'the morality of authority', which is the tendency to endorse their parents' judgments of their behaviour and develop the desire to act in accordance with these judgments (Rawls 1971, §70). As children start participating in associations, such as the school or the neighbourhood, they adopt the 'morality of association'; this means that they are influenced by the moral standards and convictions of those in their association and develop the desire to conform to these standards (Rawls 1971, §71). Finally, children develop their sense of justice sufficiently when they develop the 'morality of principles', that is, when they develop the desire to comply with the requirements of justice which bind them to strangers, by recognising that they themselves benefit from living in a just society (Rawls 1971, §72).

Although this account of how children develop the sense of justice in the well-ordered society is plausible, it is incomplete, as it lacks an account of what kinds of judgments, standards, and convictions parents and associations should instill in children for them to become reasonable and rational. This is what I do in the remainder of the chapter. More specifically, I first focus on policies, such as education, that are required for states to discharge their duties towards children. Subsequently, I examine the duties that parents have towards their children and the ways in which states must ensure that parents fulfil these duties. Of course, these policies are not meant to be exhaustive, but merely an analysis of certain key examples of policies and attitudes that ought to be promoted.

Now, the distinction between state and parental duties may be challenged, even though Rawls's own account of children's upbringing in the well-ordered society relies on it too. For instance, Plato's suggestion that the family be abolished and that children be raised by communities would suggest a greater sphere of state duties than I do. I do not have the

space to examine this suggestion but it does not seem to me that it undermines my argument in any important ways. The policy implications examined in this section are meant to be viewed as a set of recommendations designed for the kinds of liberal democracies that we observe in our world. In these societies, the desire to have and rear children is experienced by *parents* and not in more *collective* forms. For this reason, we may suppose that parents bear greater moral responsibility and therefore that they should bear greater costs too, assuming a background just distribution.[25] In any case, even if we did live in radically different societies, the nub of my argument would not be affected: the duties that arise from children's interest in developing the two moral powers would still be the same, even if they were distributed differently. For instance, perhaps states would have full and primary responsibility for covering children's nutritional needs.[26] Although this would make a difference for states and parents, the outcome would still have to be that children develop the two moral powers; in this sense, indeterminacy about the distribution of duties that I describe below would not undermine my argument that a central aim of policies ought to be the satisfaction of children's interest in developing their moral powers.

2.4.1 States

We commonly accept that states have certain distinctive obligations to children, such as the obligations to educate them and offer them social and emotional goods on an unconditional basis, which are typically different from states' obligations to adults. In particular, 'they are paternalistic in nature because we feel bound to fulfil them whether

[25] I discuss moral responsibility and parenthood in more detail in Section 5.4.2.

[26] As I argue in the chapter on cognitive disability, given the structure of our societies, parents' private decision to create and raise children can be viewed as the pursuit of an expensive conception of the good. Assuming that justice should be ambition-sensitive, this implies that more duties towards children should fall on parents than states. It remains true, however, that in a differently structured society (e.g. one in which citizens would collectively decide to create children), there are reasons to think that the distribution of duties would be different.

the children in question consent to be protected, nurtured, educated and disciplined. Indeed, we think of children as people who have to be raised whether they like it or not' (Schapiro 1999, 716). According to my account of why children matter in Rawlsian contractualism, these obligations are distinctive due to their justification; that is, they arise from individuals' interest in developing the two moral powers to the requisite degree.[27]

2.4.1.1 The Epistemic Aspect of Reasonableness

The first type of states' obligations to children springs from children's interest in developing the epistemic aspect of the two moral powers. With respect to their reasonableness, this means that they should develop the capacity to understand the requirements of justice. To begin with, this implies that education ought to target children's capacity to develop internally consistent arguments, evaluate others' arguments in terms of both validity and soundness, aspire to hold views and values that are consistent with each other, and appreciate evidence (Clayton 2006, 146-7). This capacity is necessary for the development of the ability to assess arguments and to therefore understand which arguments are justifiable to other reasonable citizens. This could be achieved by coupling children's mathematical and scientific education with lessons on ethics, political philosophy, and logic, with an emphasis on critical thinking and evaluation rather than just on the analysis of different arguments. These skills can be further developed through practical courses and activities, such as holding class debates.

[27] Other theorists seem to share some version of this view. For example, Debra Satz argues that 'the idea of educational adequacy should be understood with reference to the idea of equal citizenship. Education has long been recognized as a "foundation of good citizenship," a necessary condition for full and equal membership in the political community. Education is essential to the effective exercise of political rights. As the Court reminded us in [the Superme Court case *Brown vs Board of Education*], education is required for the "performance of our most basic public responsibilities," and its denial effectively shuts out individuals from participation in society as citizens' (Satz 2007, 635) and that 'we can derive, in general terms, the nature and content of educational adequacy from the requirements for full membership and inclusion in a democratic society of equal citizens' (Satz 2007, 636).

Apart from developing these cognitive skills, acquiring a sense of justice further requires promoting certain values and attitudes, such as impartiality, civility, and reciprocity. These values and attitudes can be cultivated through group interactions; for instance, learning to cooperate with classmates in group activities can engender in children a sense of fairness and respect (Satz 2007, 637). More specifically, learning to live up to these values first requires gaining skills that pertain to deliberation and cooperation, such as 'skills related to articulating a position and the reasons for its affirmation; listening skills; the ability charitably to understand the views of others; analytical skills that facilitate a critical assessment of different positions; an appreciation of the benefits of exchanging ideas; and a commitment to reason rather than to employ attractive slogans or rhetoric' (Clayton 2006, 147). Moreover, children should learn about the content of the theory of justice that determines the distribution of the benefits and burdens of social cooperation in their society, the liberal principle of legitimacy, as well as the duties that they will come to have qua citizens. For instance, in order to understand their duty of civility and develop that attitude, they must learn about the difference between political and comprehensive reasons. Similarly, it is crucial for them to understand the merits of public reason and the reasons why they have a moral duty to refrain from appealing to comprehensive doctrines in their public deliberation (Clayton 2006, 147-8). Thus, children's current education on the value of equality, which tends to focus on issues such as gender and racial equality, should be enlarged to include broader questions about what constitutes equal respect (Clayton 2006, 150).

2.4.1.2 The Epistemic Aspect of Rationality

The second epistemic moral power that states must target through education is children's rationality – namely, their capacity to form and revise a conception of the good. Having this capacity implies that one can make an informed judgment on different comprehensive doctrines and that the conception of the good one holds at any point can be viewed as that person's own considered choice. This is what it means, after all, to have the capacity to *form* and *revise* – rather than simply *follow* – a conception of the good. For example, someone who holds a comprehensive doctrine simply because she has been brainwashed is not autonomous in the political sense, as she has not formed that doctrine and most likely lacks the capacity to revise it. This shows that developing the rational capacity requires developing the ability to think critically about different comprehensive doctrines (Andersson 2011, 295). This requirement reinforces the need for an education that targets one's reasoning faculties, such as one's argumentative skills and the capacity to evaluate different arguments.

Yet simply giving children the reasoning skills that are required for the assessment of comprehensive doctrines is not sufficient; the development of epistemic rationality further requires teaching them about the *content* of different reasonable comprehensive doctrines and encouraging them to actively reflect on that content. This can be achieved, for example, by offering religious education classes that discuss different religious beliefs, and philosophy classes that bring up different conceptions of well-being.

The question that arises at this point is whether epistemic rationality can be developed appropriately if states, schools, and public officials, including teachers, take a stance on different conceptions of the good. In other words, is a somewhat perfectionist education permissible within the constraints of a political liberal state? It might be argued that the

commitment to fostering children's epistemic rationality does not preclude a perfectionist education, as long as this does not prevent children from learning about different comprehensive doctrines and from developing their reasoning faculties. For example, it might be argued that a teacher's promoting a specific comprehensive doctrine does not prevent students from rejecting or revising it.

However, there are two problems with this approach. First, a perfectionist education might put a strain on children's ability to develop their rationality, even though it remains true that a *somewhat* perfectionist education can be consistent with the *sufficient* development of rationality.[28] The reason why there needs to be a balance, however, and why a perfectionist education can have a negative impact on the development of rationality is that, as we've seen in Rawls's account of the 'morality of authority' and the 'morality of association', children tend to look up to figures of authority and it seems likely that their capacity to critically assess different comprehensive doctrines will be restricted by an emotional need to endorse their teacher's preferred comprehensive doctrine. Similarly, they might experience internalised guilt at the thought of rejecting that comprehensive doctrine (Clayton 2006, 107).

There is, moreover, a second problem with a perfectionist education: if states or public officials take a stance on comprehensive doctrines with regards to child-centred policies, they will implicitly evaluate the lives that adults lead as well. For example, if a state's education implies that a comprehensive doctrine is somehow superior to others, then this judgment will transfer to public discourse more generally. Consequently, although certain

[28] For instance, borrowing an altered version of Clayton's example, a teacher who openly endorses a specific religion can still instill rationality in children if they enable children to 'become familiarised with other religious and irreligious traditions' (Clayton 2006, 110).

anti-perfectionist theorists endorse a perfectionist account of upbringing (see Fowler 2014), I contend that political liberal states ought to refrain from endorsing a perfectionist education.[29] As we have seen, this is both because a perfectionist education is likely to hinder children's development of epistemic rationality and because it would implicitly pass a judgment on citizens' comprehensive doctrines.

2.4.1.3 The Motivational Aspect of Reasonableness

Having examined the implications of children's interest in developing the epistemic aspect of their moral powers for child-centred policies, I now turn to their interest in developing the motivational aspect of reasonableness and rationality. The questions that are relevant here are the following. What policies should political liberal states follow in order to ensure that children develop the motivational aspect of their reasonableness? In other words, what policies can instill in children the desire to comply with the requirements of justice? And what child-centred policies can engender the motivational aspect of children's rationality? That is to say, how can states facilitate the development of children's motivation to pursue a conception of the good once they have already developed the epistemic capacity to form and revise that conception?

With regards to the first question – targeting the motivational aspect of reasonableness – the main goal that ought to be included in education systems is the development of emotions that will motivate children to comply with the requirements of justice, once they have developed the capacity to understand these requirements. Merely understanding these requirements is not sufficient for the adequate development of children's *motivational*

[29] Although I assume that education is provided by the state, it should be noted that insofar as children are home-educated, parents act in *loco politicus*, not in a private capacity. Thus, the restrictions that apply to teachers would apply to them too.

reasonableness. This is because individuals often tend to act on the basis of their emotions and these emotions might sometimes prevent them from desiring to comply with the requirements of justice. Indeed, history contains various examples of unjust exclusion and discrimination which seemed to be entirely unsupported by the reasons that were recognised by those who supported those injustices, even if they did understand those reasons. For example,

> denials of equal opportunity to Jews, African Americans, Asian Americans, and (in India) members of lower castes were not supported by a clear-eyed assessment of economic costs and benefits; indeed, they resulted in a colossal waste of human capital. And as John Stuart Mill observed, the most ubiquitous and enduring exclusion of all, the exclusion of women from employment opportunities and political participation, is a bizarre policy for a utility-maximising society, and one that could be held in place only by irrational prejudice (Nussbaum 2013, 162-3).

According to Rawls, growing up in a well-ordered society regulated by a theory of justice from the family of liberal conceptions of justice is normally expected to be sufficient for the development of the effective desire to act from the principles of justice (Rawls 1971, 404). The thought is that the principles of justice ensure significant benefits for every member of society, including the least well-off, and that, for this reason, all members would be expected to develop the desire to comply with these principles. Moreover, Rawls argues, a theory such as Justice as Fairness is expected to have a more positive impact on children's sense of justice than alternative theories of justice would (Rawls 1971, §76). In order to advance this argument, Rawls focuses on a comparison with utilitarianism, claiming that the priority of the liberty principle and the difference principle show unconditional care for each *individual's* well-being, which utilitarianism lacks (Rawls 1971, 437; Munoz-Dardé and Sinclair, forthcoming). For these reasons, Rawls concludes, children that grow up in a well-ordered society are expected to become sufficiently reasonable citizens.

Although it seems plausible that growing up under just institutions is expected to have a positive impact on the development of a sense of justice, I suggest that just institutions, such as schools, ought to follow more specific guidelines on children's education and upbringing. For instance, in order to promote the development of the motivational aspect of reasonableness, schools should not merely teach their students about the *content* of the values that characterise reasonableness. Rather, they should also support practice-based learning, by encouraging children to *practise* the desired values and attitudes in and outside the classroom. For instance, in order to cultivate respect for diversity, schools could encourage 'intergroup knowledge, social integration, accommodation, and understanding' (Satz 2007, 637). To the extent that this already happens, it is justified as a policy that promotes equality of opportunity. For example, by encouraging social integration, disadvantaged children's marginalisation is addressed, leading to greater access to opportunities. While such justifications for the kind of practice-based learning I advocate are indeed important, it is interesting to note that my analysis of the state's duties to children points to a further reason in support of these policies: that is, their positive effects for the development of children's motivational reasonableness.

Since an important part of the motivational aspect of reasonableness involves experiencing the appropriate emotions, political liberal states must accept that they ought to cultivate those emotions. Preparing for this task requires conducting extensive studies on human psychology and on the behaviours that are caused by emotions such as fear, disgust, shame, anger, compassion and love. In certain cases, finding the appropriate balance may be particularly challenging. For example, as Rawls notes, it is clear that envy is at least *pro tanto* undesirable, as it is 'collectively disadvantageous': for it is 'the propensity to view

with hostility the greater good of others even though their being more fortunate than we are does not detract from our advantages' (Rawls 1971,465-6).[30] On the other hand, there are cases of benign envy, given that envy often motivates individuals to pursue similarly advantageous goods and activities for themselves (Rawls 1971, 467; Nussbaum 2013, 123). However, benign envy can, 'under certain conditions of defeat and sense of failure' develop into envy proper (Rawls 1971, 467). It is clear, then, that balancing emotions in the right way is more challenging than it seems at first. We can draw, however, on history and human psychology to draw certain preliminary suggestions.

For instance, children should be encouraged to develop a sense of extended sympathy that is not limited to specific groups. One way in which this could be achieved is through artistic endeavours, such as plays and songs, in schools, which can increase children's empathy. For example, Martha Nussbaum suggests that students participate in plays about the Birmingham bus boycott or about Mildred Jeter Loving, who 'refused to accept Virginia's criminalisation of her interracial marriage' (Nussbaum 2013, 126). Similarly, students could be asked to write essays, stories, or poems that reflect the point of view of victims of injustice.[31] Less directly, states can also act in certain symbolic ways, as when the U.S. Congress made Marten Luther King Jr.'s birthday a national holiday and encouraged children, to celebrate (Nussbaum 2013, 132).

It is interesting to note that if this account of the function of political emotions as part of individuals' motivational reasonableness is correct, then states may be morally required to fund artistic projects and to promote the ones that elicit the appropriate political emotions.

[30] For a discussion of envy, see Rawls 1971, 464—474.
[31] For evidence on the compassion-eliciting effects of imagining a person's suffering, see Batson 1991 and Batson 2009.

This seems to be a curious implication of Rawlsian contractualism, given that anti-perfectionist theories are typically viewed as inconsistent with this requirement. At the very least, however, this requirement arises with regards to children[32] who have not yet developed the two moral powers; therefore, funding and promoting certain artistic projects whose target audience includes children may be among a political liberal state's morally required policies.

Now someone might object that a state that influences its future citizens' emotions and *motivational* reasonableness cannot be legitimate. The thought here is that legitimacy requires citizens' consent and, by manipulating citizens' emotions, states manipulate their citizens into consenting to the exercise of political power. That is, states effectively create citizens who cannot possibly object to them. These states cannot be viewed as legitimate, the objection goes, because legitimacy requires free, informed, and unmanipulated consent.

However, we may doubt whether one's consent is valuable if it does not meet certain conditions, such as the possession of reasonableness and rationality. Indeed, this is why we do not typically view children as agents whose consent bears moral significance; for they have not yet developed the two capacities that validate one's consent. Thus, even if we accepted that it is better if we develop our sense of justice on our own, we would not have to accept that 'it is better for us to have unmanipulated political convictions *tout court*' (Clayton 2006, 143). Enabling children to develop their moral powers, even if it involves influencing their emotions, does not, therefore, undermine a state's legitimacy.

[32] I will consider whether cultivating political emotions could ever be permissible with regards to adults in the chapter on unreasonable persons.

We thus reach the conclusion that states have a duty to enable children to develop the motivational aspect of reasonableness – that is, the desire to comply with the requirements of justice – by presenting children with reasons for their compliance, by cultivating their capacity to assess these reasons, and by fostering the appropriate emotions. As we have seen, states can contribute to children's development of motivational reasonableness through a number of child-centred policies, from including certain activities and courses such as role-playing and moral philosophy in the national curriculum, to funding and promoting artistic projects that elicit the required political emotions.

2.4.1.4 The Motivational Aspect of Rationality

Having examined the implications of states' duty to enable children to develop their motivational reasonableness, I now turn to the issue of motivational rationality: what child-centred policies can engender the motivation to pursue a conception of the good in those who already have the epistemic power to rationally hold such a conception? Clearly, our discussion of education is not complete before we examine how education can facilitate the development of the motivational aspect of children's rationality. By drawing on certain examples of individuals who cannot bring themselves to pursue their conception of the good despite having formed one due to addiction or lack of self-confidence, we observe that there are certain values and psychological dispositions that education must promote: self-love, self-respect, self-confidence, determination, and perseverance. Once again, modifying the state curriculum and emphasising the significance of creative activities can serve to promote these values and dispositions. Given, however, that these traits are often developed through children's relationships and interactions with others, as we see in Rawls's account of the 'morality of authority' and the 'morality of association', it is crucial that teachers promote their development through their behaviour towards both children and

adults when acting qua teachers.[33] One way of achieving this is by requiring teachers to attend certain psychology classes that emphasise the symbolic significance and the possible implications of their behaviour and attitudes. In a non-ideal society such as ours, this behaviour may need to vary across different groups to redress issues such as lack of self-confidence experienced by specific groups due to historical injustice and exclusion.

The duty to promote the development of the motivational aspect of rationality in children seems to produce a further reason in favour of a neutralist education with regards to different reasonable conceptions of the good. As we have briefly seen, this is because children are less likely to be motivated to follow their own conception of the good if they know that a figure of authority in their lives, such as the state, a teacher or, as we will see in more detail below, their parents, encourages them to endorse a different conception of the good; for opposing an authority figure's conception of the good is likely to generate feelings of guilt, shame, and low self-esteem (Clayton 2006, 107; Green 1995, 264-7). Given that, as we have seen, a perfectionist education would also not be publicly justifiable, we can conclude that children's education should not take a stand on different reasonable comprehensive doctrines; instead, it should focus on promoting all-purpose values and emotions that can increase children's motivation to pursue their own reasonable conception of the good, whatever that may be.

2.4.2 Parents

2.4.2.1 The Epistemic Aspect of Rationality

[33] Of course, this does not mean that education is sufficient for the development of these values and psychological dispositions. As we will see in the next section, parenting is similarly if not more important for the development of motivational rationality.

Having seen the implications of the interest in developing the two moral powers for states' duties towards children, I now turn to the question of parents' duties. What kind of parenting is required in order to satisfy children's interest in developing reasonableness and rationality? And how can states encourage this kind of parenting through public policy? These are the main two questions that I focus on in this section.

First, let us consider the issue of *epistemic* rationality: that is, the capacity to form and revise a conception of the good. We have seen that the development of this capacity requires the development of certain cognitive skills and teaching children about a large number of reasonable conceptions of the good. It would be overly demanding, however, to expect parents to have the ability to further these skills or the knowledge that is necessary in order to teach children about comprehensive doctrines, even though parenting lessons could improve this ability. Consequently, the development of the *epistemic* aspect of rationality is best promoted through state-led policies such as the ones we saw in the previous section. In fact, this is precisely why mandatory education is justified: since teaching children requires difficult and time-consuming training that most parents cannot receive, the burden of ensuring the development of children's epistemic rationality falls on teachers, who have acquired the requisite knowledge.

The question that arises at this point is whether parents can permissibly pass on their own conceptions of the good on their children. In other words, would somewhat perfectionist parenting be permissible? It might seem that a commitment to political liberalism implies a commitment to an anti-perfectionist, political liberal upbringing, just as it implies a commitment to an anti-perfectionist education. Indeed, Matthew Clayton has influentially argued that the public reason restriction should apply to children's upbringing as well; that

is, that children's upbringing should 'be guided by ideals and principles that do not rest on the validity of any particular reasonable comprehensive doctrine' (Clayton 2006, 95). In a nutshell, the thought is that the exercise of political power shares important similarities with the exercise of parental power. And given that the family is one of the core institutions of the basic structure, the liberal principle of legitimacy ought to apply to parental power as well. This means that parental power ought to be justifiable to reasonable and rational persons. And given that reasonable and rational persons may come to reject the inculcation of specific comprehensive doctrines in their childhood, they would not find perfectionist parenting to be justifiable. Thus, Clayton concludes, the requirements of public reason ought to apply to parental power as well (Clayton 2006, Ch. 3).

As we saw in the previous section, it is indeed true that there is a possibility of future reasonable rejection of a perfectionist upbringing. This is because such an upbringing may cause feelings of guilt and shame that negatively affect one's capacity to form, revise, and pursue a conception of the good. Thus, we see that there is a reason in favour of anti-perfectionist parenting.

However, I suggest that this is only one consideration that should be balanced against others. A second consideration that pulls to a different direction is that a somewhat perfectionist education can be good for both children and parents because both have an interest in forming close, intimate relationships with each other (Brighouse and Swift 2014; Fowler 2015). This interest need not be justified in comprehensive terms; the thought here is that having close, intimate relationships with one's parents can enable one to build the kind of self-respect that is required for epistemic rationality. Similarly, parents have an interest in having close, intimate relationships with their children, given that this is part of

what it means to have the desire to have and rear children. Yet forming these relationships requires a sphere of freedom: in order to feel connected to their children, many parents feel the need to share their core values and beliefs with them. This might be the case for a number of reasons. For example, if a parent holds a strong belief in a specific religion, she is likely to believe that she would be compromising her child's well-being if she did not instill that belief in the child. A full commitment to one's comprehensive doctrine often requires passing it on to one's children. And even if this is not the case, the way in which humans form intimate emotional bonds with others often requires the freedom to communicate one's ideas and values. If parents are not free to do so with their children, the argument goes, they might not be able to feel as close to them as they would otherwise do. This is not just bad for parents; it is bad for children as well, given that they have an interest in forming these intimate relationships because the nature and quality of these relationships affect the development of the two moral powers. So, we see that reasonable and rational persons would not *only* have reasons to oppose perfectionist parenting; they would also have an opposing reason in favour of such parenting.

Balancing these reasons suggests that a *somewhat* perfectionist parenting may be permissible, as long as it does not interfere with the *adequate* development of a child's rationality. For instance, borrowing Clayton's example, 'parents who take their children to church to enable them to experience the traditions and practices of the church, who also allow their children genuinely to become familiarised with other religious and irreligious traditions, may not violate their children's autonomy' (Clayton 2006, 110). Parents who have themselves grown up in a well-ordered society are expected to recognise the merits of public reason and, by extension, the significance of enabling their children to develop their rationality. In all other cases, parenting lessons would teach parents about the merits

and the practical ways of finding a balance between their desire to share their comprehensive doctrines with their children and the requirements that spring from their children's interest in developing their epistemic (and motivational) rationality.

2.4.2.2 The Epistemic Aspect of Reasonableness

Let us now turn to the parental duties that arise from children's interest in developing epistemic reasonableness: that is, the capacity to understand the requirements of justice. As in the case of epistemic rationality, it would be overly demanding to expect parents to teach their children the cognitive skills that might be required in order to understand a theory of justice. There is, however, a range of parenting techniques that affect children's capacity to engage in moral reasoning and which parents can realistically engage in if they are provided minimal training.

According to the literature, the main factor that influences a child's capacity to understand the requirements of justice is a parent's attitude towards communication and discipline. For example, 'authoritative' parents, who are 'loving, controlling, communicative and set high maturity demands for their children' tend to have children with higher moral functioning (Berkowitz and Grych 1998, 379). By contrast, 'authoritarian' parents, who are highly controlling and demanding but affectively cold, hostile, and uncommunicative', as well as 'permissive' parents, who are 'loving and communicative but wield little control and set few demands for mature behaviour' hinder their children's development of moral reasoning (Berkowitz and Grych 1998, 379).

One of the main techniques that authoritative parents use, and the most crucial one for the development of epistemic reasonableness, is induction. This means that parents ought to

communicate with their children by giving them reasons for their actions or inactions. This is important for two reasons. First, it is important for parents to lead by example by showing respect both to their children and to others. Given that respect requires among other things the capacity to satisfy the duty of civility by giving justifications to co-citizens, showing respect to children requires giving reasons to them. The second reason why this is important is because it teaches children *how* to give reasons. This is important because, together with learning about the importance of giving reasons, children also need to learn the practical ways of doing so.

There are a number of techniques that parents can use in order to foster their children's development of epistemic reasonableness through the use of induction. For example, to the extent that it is possible, parents should explain the concepts of justice and fairness to their children by drawing on examples from their everyday life, such as why teachers ought to be impartial or why children themselves ought to treat others with respect. In these and other dialogues, apart from *explaining* concepts, parents can also use Socratic questioning to teach their children to give reasons (Walker & Taylor 1991). Offering explanations for ideas, actions, or inactions, and engaging in Socratic questioning can further children's capacity to give reasons, thereby furthering their epistemic reasonableness. As in the case of teaching parents to foster their children's epistemic rationality, parenting lessons can teach parents the value of being authoritative and of encouraging open and supportive communication within the family.

2.4.2.3 The Motivational Aspect of Rationality

Turning to the interest in developing the motivational aspect of the two moral powers, let us first focus on rationality. What parental duties arise from children's interest in developing the capacity to *pursue* their conception of the good?

The literature suggests that the main determinant of children's development of motivational rationality is the love they receive from their parents. In defining what it means for a parent to love her child, I follow Matthew Liao, whose definition captures the main relevant points:

> to love a child is to seek a highly intense interaction with the child, where one values the child for the child's sake, seeks to bring about and maintain physical and psychological proximity with the child, seeks to promote the child's well-being *for the child's sake*, and desires that the child reciprocate or, at least, respond to, one's love (Liao 2015, 76, my emphasis).

This definition highlights that loving someone is not merely a behavioural or attitudinal matter. Rather, it contains an emotional aspect: children have a strong interest in their parents *feeling* love for them. The behavioural and attitudinal aspects, however, should not be ignored either: parenting must be shaped by the love parents feel for their child as opposed to, for example, by the parent's own fears or projections. This means that it is not sufficient for parents to love their children; children must *receive* this love.

There are at least four reasons why children have an interest in their parents' love, which further determine *how* parental love must be given (Liao 2015). First, children have an interest in developing trust in others and parental love enables a child to develop that capacity. This interest is connected to children's interest in developing motivational rationality because most conceptions of the good require the capacity to trust others. For

instance, many conceptions of the good involve intimate relationships with others. If a child does not develop the capacity to trust others, she will not have the capacity to form or maintain intimate relationships (Erikson 1950; Bowlby 1980). For example, an adult who cannot trust others is likely to experience anxious attachment to others, which might result in abusive relationships (Rodriquez et al. 2015). Moreover, the capacity to trust others is necessary for a number of all purpose goods that enable us to pursue most, if not all, conceptions of the good. According to Erikson and Bowlby, children develop the capacity to trust others when they receive love from their parents, when their parents attachment to them is secure, stable, and when their parents respond predictably and reliably to their children's needs, particularly in the care they receive in the first year of their life (Erikson 1950; Bowlby 1980; Liao 2015).

The second reason why children have an interest in their parents' love is that parental love is necessary for one to develop a positive sense of self (Rawls 1971, 406–7; Liao 2015, 79). Our sense of self is essentially our belief about our own capacity to succeed or fail in our pursuit of our conception of the good. Having a negative sense of self implies that one will often refrain from pursuing one's conception of the good because she will believe that doing so would be fruitless, even if this does not in fact reflect the truth (Berkowitz and Grych 1998, 384). The capacity to pursue one's conception of the good then requires a positive sense of self and in order to develop that sense of self, children need to receive parental love (Pringle 1986, 35). This means that parents ought to believe in their children, albeit in a realistic manner, which further implies that they should accept their children for who they are. They should set reasonably high expectations, and encourage their children to reach their own potential and to build both the desire to do so and the belief that they have the capacity to achieve this.

The third reason why children have an interest in being loved by their parents is that receiving parental love teaches them how to love. As we have seen, knowledge of how to love is crucial for one's motivational rationality because the capacity to form intimate, loving relationships is a crucial part of many conceptions of the good. If children do not develop the capacity to love, they will never be capable of pursuing these conceptions of the good. But in order 'to know how to love, a child needs to be loved, because love is a complex phenomenon that one can only learn by having received it and by having the opportunity to practice it on someone' (Liao 2015, 81).

Finally, love provides children with the motivation to accept their parents' discipline, which is necessary for them to develop a number of capacities that are all-purpose goods (Liao 2015, 82-3). The thought here is that good kinds of discipline enable children to remain safe, to explore the world, to learn, and to act in ways that are good for them, such as learning to eat healthily or to not abandon their activities on a whim or when they encounter a few difficulties. According to studies, receiving love is necessary for children's acceptance of parental discipline and, by extension, for the development of these attitudes and capacities (Parke and Walters 1967; Liao 2015). This is both because children are more likely to accept discipline from someone that they know loves them, and because loving parents tend to make reasonable demands and follow authoritative discipline patterns (Liao 2015, 83).

In fact, studies confirm that when children do *not* receive parental love, as for example in the case of institutionalised children, they tend to

> become ill more frequently, their learning capacities deteriorate significantly, they become decreasingly interested in their environment, they fail to thrive physically by not gaining weight or height or both, they suffer insomnia, they are constantly depressed,[34] and they eventually develop severe learning disabilities (Liao 2015, 88).

We can therefore conclude that overall, children have a very strong interest in receiving parental love, both because this enables them to develop important capacities and because the denial of that love harms them in a number of ways. Given that parental love is connected with the development of children's motivational rationality, parents have a duty to do as much as possible in order to love their children. Notice that my argument does not require accepting that there is a duty to love one's child for, as Brighouse and Swift have argued regarding the ethics of parent-child relationships (Brighouse and Swift 2014), we cannot have a duty to do what we cannot do,[35] and there is disagreement on whether we can control our love.[36] However, because the development of a child's motivational rationality depends on the love a child receives, parents have a duty to do whatever they can in order to develop love for their children, and they certainly have a duty to adopt the attitudinal and behavioural aspects of love by following, for example, an authoritative parenting style.

2.4.2.4 The Motivational Aspect of Reasonableness

Having examined the implications of children's interest in developing motivational rationality for the duties that their parents have, I now turn to children's interest in developing motivational reasonableness, i.e. the willingness to comply with the

[34] For instance, we have seen that parental love is necessary for the development of self-esteem and a lack of self-esteem is linked to mental health issues such as depression and anxiety (Harter 1997).
[35] For a broader discussion of whether ought implies can, see, for instance, Kant 1971/1797; Mill 1979/1861; Ross 1930; Dahl 1974; Sinnott-Armstrong 1984.
[36] It should be noted, of course, that if it *is* possible for us to control the love we feel for others, my argument implies that parents do have a duty to love their children.

requirements of justice. First, we should note that the interest in being loved is relevant here too. Given that low self-esteem and a negative conception of self can lead to addictions which can then lead to the violation of duties of justice, receiving parental love is necessary for one to develop motivational reasonableness. Importantly, parenting can further affect the development of a very large number of pro-social attitudes and emotions.

First, parenting greatly influences a child's social orientation: that is, the desire to engage in social interaction and to form and maintain relationships. As we see in Rawls's writings, children's capacity to develop the 'morality of association' depends on their tendency to develop friendly feelings towards associates (Rawls 1971, 429-30). The literature on this tendency suggests that one's social orientation is a crucial part of this tendency (and therefore of reasonableness) because a 'pervasive detachment from and disinterest in social relationships' leads to 'an antisocial lifestyle in which the rights of others are ignored or violated' (Berkowitz and Grych 1998, 373; Weiner 1980). The reason why parenting affects one's social orientation is that the development of that orientation depends on whether one learns to form secure attachments in the first years of one's life (Berkowitz and Grych 1998, 373; Park and Waters 1989). In order to enable their children to develop secure attachments with them, parents ought to be 'attuned to their infant's signals' (Ainsworth et al. 1978) and 'respond to them consistently and rapidly' (Ainsworth et al. 1972). Moreover, they ought to engage in consistently loving behaviour by coming into affectionate physical contact with their children (Clarke-Stewart 1973), by refraining from being hostile or angry (Main et al. 1979), and by being diligent in their care of their child (Egeland and Sroufe 1981).

The second pro-social attitude that is influenced by parenting is a child's self-control. The reason that self-control is associated with one's motivational reasonableness is that

> developmental implications of early differences in inhibitory control include differences in multiple aspects of morally relevant conduct, such as the tendency to violate prohibitions while without surveillance ..., adolescent drug use ..., self-control in the face of temptation ..., or empathy (Kochanska et al. 1997, 264).

Parents ought to help their children develop self-control by modeling self-control themselves; by explaining to them that actions lead to certain consequences; by ensuring that the household contains minimal or no conflict; and by setting high but realistic expectations for them (Berkowitz and Grych 1998, 375).

Third, parents can affect the development of a child's motivational reasonableness by affecting their child's empathy. Given that lower levels of empathy have been linked with antisocial behaviour, the development of empathy is crucial for ensuring that one is willing to comply with the requirements of justice (Gibbs 1987). In order to positively affect the development of their child's empathy, parents ought to engage in induction, offer their child reasons and justify their actions, and avoid power assertion and love withdrawal (Hoffman 1983; Berkowitz and Grych 1998, 377).

The fourth trait that affects a child's motivational reasonableness and that is influenced by parenting is the development of conscience. Conscience affects our capacity to comply with the requirements of justice in two ways. First, it enables us to have appropriate emotions in response to transgression. For example, it leads us to experience 'guilt, apology, empathy for the victim, etc.', which prevent us from violating our duties of justice (Berkowitz and Grych 1998, 377). Second, it encapsulates our capacity to engage in 'moral

regulation or vigilance, which encompasses the classical internalisation of standards along with confession, reparation and monitoring of others' wrongdoing' (Berkowitz and Grych 1998, 377). According to research on the development of conscience, parents ought to use induction and avoid power assertion in order to ensure that their children internalize moral standards and the capacity to feel guilt (Hoffman and Saltzstein 1967; Allinsmith and Greening 1955).

Finally, parenting influences a child's capacity to engage in altruistic behaviour, which has also been linked to the willingness to comply with the requirements of justice. This is because altruistic children tend to be 'active, sociable, competent, assertive, advanced in role taking and moral judgment, and sympathetic' (Eisenberg and Mussen 1989, 151). According to Eisenber and Mussen, in order to facilitate their child's capacity to behave altruistically, parents ought to be 'nurturant and supportive, model altruism, highlight the effects of actions on others, use induction, establish clear expectations for mature behaviour, create opportunities for their children to manifest responsibility for others', and model compassion and altruistic behavior themselves (Berkowitz and Grych 1998, 378).

I suggest that in order to enable parents to gain an understanding of their parental duties and the knowledge that is required in order to follow these parenting techniques, states ought to require all parents to attend parenting lessons in order to ensure that they learn how to discharge their parental duties, which are associated with their children's interest in developing the two moral powers.

2.5 Conclusion

In this chapter, I have explored the problem of children's exclusion in Rawlsian contractualism. I have argued that, due to its conceptions of personhood and social cooperation, Rawlsian contractualism might exclude children in three ways. First, they might be excluded from the constituency to whom justification is owed; second, their interests might be excluded from justice as fairness; lastly, there are good reasons to suppose that the principles of justice do not apply to them. As it stands, Rawls's theory of justice and legitimacy faces these urgent problems.

The main purpose of this chapter was to resolve these problems. In the third section, I argued that all beings who have the potential to develop the two moral powers, including children, have an interest in doing so. By incorporating this interest into Rawls's theory of justice and legitimacy, I have argued that the theory can offer satisfying responses to the question of children's exclusion.

Finally, in the last section of this chapter, I have examined the implications of my argument for a Rawlsian state's child-centred policies. More specifically, I have proposed a range of duties that fall on states and parents that aim to ensure the development of children's epistemic and motivational reasonableness and rationality. Thus, I hope to have shown that children can be included in Rawlsian contractualism and that this can be done in a way that provides us with useful and intuitively plausible guidelines for policy-making.

Chapter 3

Foetuses

3.1 Introduction

As we have seen, public reason expresses the Rawlsian contractualist's key commitment that the exercise of political power ought to be justifiable to those who are subjected to it. As Rawls argues in *Political Liberalism*, laws on constitutional essentials and issues of basic justice ought to be justifiable to an idealised constituency of reasonable and rational citizens. This view finds expression in the liberal principle of legitimacy, according to which

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

However, we would expect that even an idealised constituency of reasonable citizens would be characterised by reasonable pluralism on comprehensive issues,[37] such as religious beliefs or views on what makes life valuable. We have seen that this is due to the burdens of judgment – sources of disagreement, such as the tendency that different

[37] As we have seen, a moral conception 'is comprehensive when it includes conceptions of what is of value in human life, and ideals of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole. A conception is fully comprehensive if it covers all recognized values and virtues within one rather precisely articulated system; whereas a conception is only partially comprehensive when it comprises a number of, but by no means all, non-political values and virtues and is rather loosely articulated. Many religious and philosophical doctrines aspire to be both general and comprehensive' (Rawls 2005, 13).

individuals have to weigh competing considerations differently – which obtain under the liberal institutions of a well-ordered society (Rawls 2005, 55-57).

If reasonable disagreement on comprehensive doctrines is indeed inevitable under liberal institutions due to the burdens of judgment, it follows that the justification of political power cannot entail an appeal to specific comprehensive doctrines. For if it does, then that exercise of political power will not be justifiable to those who do not share that comprehensive doctrine. This implies that states should not explicitly or implicitly endorse any particular comprehensive doctrine on which reasonable citizens could reasonably be expected to disagree; instead, they should rely on *political* values or public reasons that all reasonable citizens can accept.

A number of critics of Rawlsian contractualism have argued that the theory is unable to give satisfying answers to the question of whether abortion ought to be permitted. Given the seriousness of this question, they claim that this inability undermines Rawlsian contractualism itself. Although different critics identify different problems, they all agree on the source of these problems: they claim that Rawlsian contractualism gives unsatisfying answers with regards to abortion because giving a satisfying answer would require appealing to a comprehensive doctrine (such as a metaphysical view of personhood); and, as we have seen, doing so is ruled illegitimate by Rawlsian public reason. I will argue in this chapter that, although it is true that the public reason restriction causes problems for the theory's implications for abortion, the *primary* reason why abortion is a difficult issue for Rawlsian contractualism is that foetuses are excluded from both the justificatory constituency and the constituency to whom the principles of justice apply. As we have seen, this is because foetuses do not meet the conditions for political personhood – that is,

the threshold of reasonableness and rationality. Since they are excluded, those included face the challenge of reaching a consensus on how they are to be treated; yet, as the critics point out, this consensus is impossible given reasonable disagreement on comprehensive doctrines.

There are two main problems that critics focus on. The first is the theory's supposed incompleteness: Rawlsian contractualism, it is argued, cannot provide us with a definite answer on whether, and under what circumstances, abortion ought to be permitted, because any definite answer would require appealing to comprehensive doctrines, such as metaphysical conceptions of personhood or conceptions of well-being.

This objection can take two forms (Gaus 1996, 151–8). First, it is argued that public reason is *indeterminate*. According to this form of the objection, determinacy on abortion requires saying *something* about prenatal moral status. Given that public reason rules out precisely these kinds of claims, it follows that it is not consistent with any arguments about the permissibility of abortion, and so it is indeterminate whether abortion is permissible under political liberalism. In fact, critics argue that public reasoners' endorsement of a pro-choice position on abortion reveals an implicit reliance on the non-public reason that the foetus is not, metaphysically speaking, a person; for, if the indeterminacy argument is true, there is no publicly justifiable way in which public reasoners could take *any* stance on abortion, including a pro-choice one (Friberg-Fernros 2010; Sandel 2005; Shaw 2011). According to the second form of the incompleteness objection, public reason is *inconclusive*: if several different positions are consistent with the theory, then it cannot provide us with decisive guidelines on policy-making (Reidy 2000, 70).

The second problem that critics have identified is a problem of counterintuitive implications: they argue that Rawlsian contractualism's implications for the issue of abortion are so counterintuitive that the theory must be rejected. As with the incompleteness objection, there are two versions of this objection. The first has been defended by Jeremy Williams. Rawlsian contractualism is, he claims, neither indeterminate nor incomplete, because it *can* appeal to a relevant public reason: namely, the reason given by women's basic rights and interests (Williams 2015, 25). What follows, however, is that Rawlsians must always endorse the permissibility of abortion, even in late-stage pregnancies (call this the *extreme permissibility argument*). This implication of Rawls's theory, Williams argues, is deeply objectionable and undermines the plausibility of the theory.

The second, contrasting interpretation of the counterintuitive implications objection is the *extreme non-permissibility argument*, according to which Rawlsian contractualism cannot possibly permit abortion, since the parties in the Original Position would choose to ban abortion out of fear that they might turn out to be among the aborted foetuses (Evers 1978, 111; Shaw 2011, 96).

In this chapter, I contend that all of these objections to Rawlsian contractualism misfire, and that the theory's implications for abortion are not problematic in the ways that critics suggest. I begin by demonstrating that the incompleteness objection, which is presented as an objection to public reason's capacity to give a determinate verdict on the permissibility of abortion is, in reality, an objection to the basic premises of public reason – *not* to its implications. Yet the aim of both variants of this objection (i.e., indeterminacy and inconclusiveness) is to show that, even if one accepts the premises of public reason, its

implications for abortion provide us with strong reasons in favour of rejecting the theory. Thus, I argue that, if one truly accepts the premises of public reason and focuses on the theory in the context of abortion, both the indeterminacy and the inconclusiveness interpretations of the argument lose their force. Second, I argue that both variants of the counterintuitive implications objection (i.e., Williams's extreme permissibility argument and Shaw's extreme non-permissibility argument) collapse once we observe that the right to abort a foetus merely implies removing it from a woman's body, as opposed to ending its existence.

The abortion-based objections to Rawlsian contractualism are therefore not compelling, because the theory's implications for abortion are determinate, not objectionably inconclusive, and not particularly objectionable. However, my response to these objections raises questions about foetuses' exclusion from the realms of justice and legitimacy in a new form. Should states and medical professionals should offer medical assistance to removed foetuses in order to enable them to live? These worries might give rise, once again, to the incompleteness objection, since Rawlsian contractualism might seem indeterminate or inconclusive on whether medical assistance should be offered. As I argue, however, in these cases the responses I provide to the charge of incompleteness in the first section safeguard Rawlsian contractualism from this additional worry.

3.2 The Incompleteness Objection

3.2.1 Indeterminacy

It is often objected that Rawlsian contractualism is indeterminate on the issue of abortion: the theory cannot say *anything* about the permissibility of abortion, because any views on abortion must necessarily rely on a view about foetuses' moral status (Friberg-Fernros

2010, 37; Sandel 2005, 133; Shaw 2011, 94). Given that Rawlsian contractualism requires that we refrain from appealing to comprehensive doctrines, including metaphysical views about personhood, some critics argue that it cannot reach *any* conclusions on the permissibility of abortion. Thus, they conclude that political liberals who defend abortion, as Rawls did, implicitly rely on the comprehensive view that the foetus is *not* a person (Friberg-Fernros 2010; Sandel 2005; Shaw 2011). Similarly, if a political liberal argued against abortion, their argument would implicitly rely on the assumption that the foetus *is* a person. Thus, those who press the indeterminacy objection highlight that all views on abortion are inconsistent with Rawlsian contractualism; the theory cannot tell us *anything* about the permissibility of abortion without violating its key commitment to public reason. And assuming that being able to reach a considered judgment on abortion is very important for any theory, the inability of Rawlsian contractualism to do so poses a very serious problem to the theory.

Clearly, the plausibility of the objection depends on whether determining the permissibility of abortion indeed requires assessing whether the foetus is a person. Those who press this objection argue that an assessment of foetuses' moral status is necessary because the key question about abortion is whether or not it constitutes murder; and in order to respond to that question, we must know whether a foetus is a person with a basic right to life. Since the 'value of human life trumps (almost) all other values', critics argue that, by endorsing abortion, political liberals implicitly assume that foetuses do not meet the conditions for personhood (Shaw 2011, 94).

At this point, a political liberal might try to defend abortion by appealing to gender equality; this would not violate the requirements of public reason, since equality is a political value.

To this, however, critics respond that we need to know who counts as a person in order to respond to the even more fundamental question of 'equality for whom?'. That is, if we care about equality, then it might seem controversial that 'the foetus is clearly not being treated as equal to the woman' and any justification for that inequality would certainly rely on the claim that foetuses are not persons, while pregnant women are (Shaw 2011, 94). Yet drawing this distinction between the moral status and, by extension, the rights, of foetuses and that of women requires, once again, relying on a metaphysical view of personhood.

However, there are a number of different yet mutually consistent responses to the indeterminacy objection that Rawlsians can draw upon. First, the assumption that abortion cannot be morally permissible if foetuses are persons has already been impugned by Judith Jarvis Thomson (Thomson 1971, 49). According to Thomson's famous violinist example, we are asked to imagine that we can only prevent a violinist from dying if we remain plugged into the violinist's circulatory system for nine months. Thomson's violinist example indicates that we may have strong reasons to let the violinist die if we are not responsible for his situation and if keeping him alive infringes on our rights or imposes disproportionately great costs upon us. In the case of abortion, the violinist analogy justifies, at the very minimum, aborting a foetus in cases of rape or when the pregnancy threatens a woman's health or life; importantly, this is true even if the foetus is a person who has a right to life. Therefore, the assumption that a woman's rights and interests can never be sufficient to justify abortion even if foetuses were persons is misguided.

Second, even if those who press the indeterminacy objection are right in arguing that Rawlsians who argue in favour of abortion implicitly assume that foetuses are *not* persons, their argument does not prove that Rawlsians do so by relying on a comprehensive doctrine.

In order to argue that foetuses lack rights it is sufficient to claim that they are not persons in the *political* sense; arguing that they cannot count as persons in a metaphysical sense is not necessary. As we have seen, in order to count as a person in the political sense, one must possess the two moral powers: reasonableness and rationality. This is because society is viewed as social cooperation with the purpose of mutual advantage and those who lack the two moral powers cannot participate in social cooperation under fair terms. Those who do not possess the two moral powers are therefore not viewed as equals in the political sense and do not, therefore, necessarily possess the same rights and liberties that political persons do.

Of course, as we have seen in the chapter on children, even those who cannot count as persons in the political sense might have certain interests that give rise to duties that apply to those who do count as persons. Indeed, the interest that children and future generations have in developing the two moral powers seems to apply to foetuses as well. Thus, it might seem that Rawlsians cannot defend a pro-choice position on abortion because doing so would disregard foetuses' (and other future persons') interest in developing the two moral powers, thereby revealing temporal partiality in favour of existing persons. However, even if states affirm foetuses' interest in developing the two moral powers, they need not accept that this interest outweighs the basic rights and interests of full persons. Given that women have a right to bodily integrity (which is the main right that the violinist analogy implicitly appeals to), and given that pregnant women are typically sufficiently reasonable and rational and therefore count as free and equal persons, while foetuses do not, women's basic rights and interests take priority over the basic rights and interests of foetuses. Thus, we see that Rawlsians can endorse a specific answer to the question of abortion without appealing to any comprehensive doctrines; the fact that foetuses do not count as persons in

the political sense shows that, even if states acknowledge their interest in developing the two moral powers, that interest cannot outweigh the basic rights of those who currently count as persons here and now.

Critics of public reason might still insist that we cannot reach a conclusion about abortion without assessing whether the foetus is a person. Their insistence on this point would reveal that they rely on a conceptual confusion, however, as they would be claiming that determinacy on the question of abortion is *impossible* in the absence of a reliance on comprehensive doctrines, while what they mean is that doing so is *objectionable*. For if it is clear that Rawlsian contractualism *can* take a stance on abortion, given that foetuses are not persons in the *political* sense and given that women's basic rights are at stake, their criticism cannot be that Rawlsian contractualism is *indeterminate*. Their objection therefore cannot be that Rawlsians are inconsistent because it is *impossible* to permit abortion without relying on a comprehensive doctrine; rather, their objection would be that Rawlsians *should not* advance this kind of pro-choice argument because they *ought to* take foetuses' moral status into consideration. The problem for those who press the indeterminacy objection in this way is that they do not justify their claim that we ought to take metaphysical personhood into consideration; they merely *assume* that we ought to do so and proceed to conclude that political liberals implicitly do so as well. This seems even more significant once we notice that it is possible to reach a conclusion about abortion without appealing to comprehensive doctrines. We thus see that the nub of the indeterminacy objection is not that public reason is indeterminate but that it is mistaken; it is an objection to the abstract and most basic commitments of political liberalism, rather than to its coherence or implications for abortion.

Although this different objection lies beyond the scope of this thesis, there are strong *prima facie* reasons to dismiss it. If a theory or certain principles are truly plausible, it might be best that we accept their implications for applied issues. Thus, if we assume that Rawlsian contractualism is indeed plausible, we might agree that we ought to avoid appealing to comprehensive doctrines such as religious and metaphysical views when arguing about abortion laws. In any case, however, as I have argued, even if Rawlsian contractualism is not in the end plausible, it is not indeterminate on issues of excluded parties such as the case of abortion.

3.2.2 Inconclusiveness

The incompleteness objection to Rawlsian contractualism is sometimes interpreted in a different way, as the charge that the theory is *inconclusive* rather than indeterminate. In the context of abortion, the charge is that Rawlsian contractualism does not give rise to a single, clear position on the matter (see Freeman 2007, 242–243; Gaus 2011, 303–333; Quong 2011, 204–212, 285–287; Rawls 1996, 240–241; Schwartzman 2004; Williams 2000). Instead, there are several different, conflicting conclusions that are consistent with the theory's main commitments. This is because the political values at stake can be defined, assessed, and prioritised in different ways. Let us assume, for example, that the values that we ought to take into consideration when discussing the issue of abortion are indeed the ones that Rawls identifies: the equality of women, the due respect for human life, and the ordered reproduction of society over time (Rawls 2005, 243). Given that these concepts will have to be given precise content and relative weight in order to determine a policy decision, many conflicting positions may be compatible with taking them to be the only values at stake. For instance, one might believe that the equality of women requires that we equalise their social disadvantages by providing them with additional resources,

employment benefits, and help with their parental duties rather than by giving them access to abortion. Unless we therefore think that equality requires eliminating *genetic* differences, such as the fact that only women can bear children, certain egalitarians might conclude that their commitment to equality does not necessarily imply that abortion is permissible. Similarly, one might argue that respect for human life implies that abortion is morally impermissible, particularly if there are other ways to achieve equality that do not entail ending a foetus's life; or one might simply think that respect for human life outweighs our other commitments, such as our commitment to equality, if and when they do clash. Lastly, we can offer various definitions of the 'ordered' reproduction of society over time: for example, we might think that we should reproduce only to the extent that individuals can live well given the scarcity of natural resources (although we might again disagree on which standard of living is good enough), that we should reproduce as much as possible, or that we should simply protect humanity from extinction. From these examples, it is already clear that, if equality, the value of human life, and society's reproduction are indeed the relevant considerations, a large number of positions on the spectrum from extreme pro-life to extreme pro-choice views may be consistent with Rawlsian contractualism, depending on the content and relative weight we give to each of these values. In this way, it seems that Rawlsian contractualism is vulnerable to the claim that it is inconclusive because it cannot give us a clear answer on the issue of abortion.

However, even if Rawlsian contractualism is indeed inconclusive, it is not clear why this poses a serious problem to the theory. Inconclusiveness is a problem that most theories face, precisely because they must give content to the normative values that they appeal to and because theorists disagree on what these abstract values require, even if they share the same commitments. This is particularly the case when applied issues are discussed, and

even more so when this occurs in non-ideal theory, since these contexts force us to consider all the details of these values and how they work together in practice. For example, luck egalitarians disagree on the best ways in which we can equalise the effects of inequalities in individuals' and groups' genetic and social luck. Consequently, a number of conflicting positions on an applied suggestion of ways of eliminating inequalities due to luck – say, the justifiability of genetic engineering as a way of eliminating genetic luck – are consistent with the theory. For instance, some might consider genetic engineering appropriate as a means of eliminating natural differences, whereas others might argue that the appropriate strategy concentrates on eliminating the social consequences of unequal natural endowments (Theofilopoulou 2015). It seems unreasonable to expect that a theory should be conclusive on every issue, particularly once we acknowledge the burdens of judgment that give rise to reasonable disagreement on our interpretation and prioritisation of political values.

Moreover, we can question whether inconclusiveness is a problem, as long as Rawlsian contractualism (or any other theory) can somehow reach a decision among competing views that would make law- and policy-making on the issue of abortion possible. As Micah Schwartzman and Andrew Williams suggest, if public reason's central arguments are valid, it would be preferable to meet practical problems with alternative decision-making solutions instead of resorting to comprehensive doctrines. We might thus endorse second-order decision-making strategies, such as 'democratic forms of mutual accommodation' or 'random decision procedures' (Schwartzman 2004, 209-14; Williams 2000, 209-211).[38] Although certain procedures, such as randomising decision-making, may not be desirable in the case of abortion, a democratic procedural solution would be possible. Assuming that

[38] For a discussion of different procedural mechanisms, see Williams 2017.

the decision-making process would include only positions consistent with the requirements of Rawlsian contractualism in the set of options given, any of these solutions would be publicly justifiable (Quong 2011, 209). Thus, even if those pressing the inconclusiveness objection rightly claim that Rawlsian contractualism is consistent with many positions on abortion, law- and policy-making remains possible.

3.3 The Extreme Permissibility Argument

According to Jeremy Williams, any appeals to prenatal moral status to defend the impermissibility of abortion will presuppose assertions either of the foetus's personhood or of the intrinsic value of human life (Williams 2015, 26). That is, we may claim that the foetus is owed respect only if we believe that it is a person with rights or if we think that all forms of human life are intrinsically valuable. Yet such claims are not permissible within the context of public reason. Given that the only publicly justifiable reasons that can be given are about women's rights, Williams concludes that Rawlsian contractualism has the counter-intuitive implication that even late term abortions are permissible.

Reconstructed, Williams's argument takes the following form:

1. Determinacy on abortion requires saying enough about either prenatal moral status or women's rights.
2. Public reason theorists cannot appeal to *any* claims regarding prenatal moral status.
3. Public reason theorists can appeal only to women's rights.
4. Public reason therefore commits us to the view that abortion is permissible regardless of the pregnancy's stage.
5. Permitting late-term abortions is deeply problematic.

C. The implications of public reason liberalism for abortion undermine the theory.

After examining and rejecting one possible way of resisting Williams's conclusion, I will press a different objection against his argument: namely, that premise 3 grounds the justifiability of *removing* a foetus and not necessarily of ending its development.

One possible objection to the extreme permissibility argument is that the first premise overlooks the possibility of appealing to the 'political value of the ordered reproduction of society' (Rawls 2005, 243). Political liberals may express a *political* worry that the full legalisation of abortion could lead to sub-optimal levels of reproduction, leading to an undesirable decline in population, for example.

However, even setting aside the empirical objections that could be raised as a response, Williams plausibly suggests that this value alone could not outweigh individuals' reproductive freedom (Williams 2015, 44). If it did, then the same value would be sufficient to restrict access to contraception when doing so would promote the orderly reproduction of society. Assuming that public reason liberals cannot appeal to prenatal moral status, they would not be able to make anything of the difference between killing a foetus and preventing the existence of a foetus (Williams 2015, 44). Unless political liberals are prepared to accept restricting access to contraception, they cannot appeal to this political value as a decisive objection to late term abortions.

It might be replied that we already *do* accept that, under certain circumstances, societal considerations outweigh individuals' liberty. For example, it is often held that, at least in the case of unjust aggression against one's country, conscription is justified. By analogy,

it may be argued that citizens' reproductive freedom may be constrained if society's reproduction is at stake. However, a number of considerations make this line of reply unattractive here. To begin with, it is not clear that conscription is indeed justified, even in the cases of unjust aggression against one's country. Since there are various arguments against conscription, we should not merely assume that the practice is justified. Given the centrality of the freedom to pursue one's conception of the good in political liberalism, Rawlsians have particularly strong reasons to question the justifiability of conscription.

Even if one accepts that conscription is justified, the analogy with procreative liberty is not clear. First, any war is a particularly urgent situation, contrary to the concern about having a falling population; the analogy would require the assumption that the danger in the latter case is that (or close to that) of imminent extinction as a polis. Second, in the case of aggression against a particular country or group, the basic rights and interests of existing people are directly at stake. By contrast, the only existing people who are negatively affected by the prospect of extinction through population decline are those who believe extinction as a polis would be so bad that it would actually affect the way in which they currently live their lives. And even in that case, those individuals' basic rights would not be threatened. Therefore, the justifiability of conscription would not suffice to show the justifiability of constraining individuals' reproductive freedom in liberal societies and, more broadly, societal considerations about the ordered reproduction of society do not look as if they outweigh individuals' reproductive freedom. Thus, Williams's conclusion that Rawlsian contractualism has objectionably radically permissive implications for abortion cannot plausibly be resisted via this strategy.

3.3.1 The Objectionable Implications of Public Reason

However, I contend that even if we accept that Rawlsian contractualism has these implications for abortion, Williams's conclusion that this is fatally objectionable does not follow. This is because women's basic rights can ground an argument in favour only of *removing* the foetus, not of killing it. Abortion is typically associated with the end of the foetus's existence because, under current regulation, which allows it only in early pregnancies, the removal of the foetus inevitably implies its death. Williams effectively assumes that late-term abortions would necessarily have the same consequences for the foetus. This, however, would be true only if there were also decisive arguments in favour of actively ending the foetus's existence. Yet political liberals' commitment to protecting women's rights cannot have that implication. Indeed, even in early-term abortions, the only justifiable intention is to remove the unwanted foetus from a woman's body, not to end its existence. Of course, some women's intention might actually be to end the foetus's development – not simply to remove it from their body. They might even attempt to justify this intention by appealing to some interest that they have, such as the interest in controlling what happens to their genetic material.[39] We may doubt, however, whether that intention could indeed be justified and whether, by extension, it could serve to justify ending the foetus's development once it has been removed. For, as we have seen in Chapter 2, a second consideration that is relevant to the question of abortion is foetuses' interest in developing the two moral powers. We saw in the previous section that the rights of current persons (such as women's right to bodily integrity) outweigh future persons' interests, which justifies the permissibility of removing a foetus; yet once a foetus has been removed, future persons' interests cannot possibly be outweighed by women's basic rights, since a woman's right to bodily integrity is thereafter irrelevant. It follows, then, that Williams's argument can at most assert only that Rawlsian contractualist implications for abortion will be radical

[39] I am grateful to Jeremy Williams for pressing me to address this objection.

in the sense of defending women's right to remove unwanted foetuses from their bodies but not in the sense that it will justify the destruction of viable foetuses so removed.

At this point, one might object that Rawlsian contractualism faces the indeterminacy objection for a different reason. This is because, it might be argued, the theory would remain indeterminate on the question of what should be done with the removed foetuses. According to this objection, the charge that Rawlsian contractualism faces is that, since it cannot reach a conclusion on the issue of unwanted foetuses, it must resort to the use of non-public reasons: any position on abortion would then have to appeal to comprehensive doctrines such as metaphysical views on whether a foetus is a person.

As I have already argued, however, the indeterminacy objection fails to explain why the considerations that public reason permits are truly insufficient in order to reach a conclusion on abortion. At worst, then, the objection against Rawls regarding abortion is that his account of public reason is *inconclusive*: that is, it does not produce a decisive answer to this question, but is instead consistent with a number of answers, depending on the context in which the question arises. Yet, as I have argued, even if public reason were truly inconclusive in this way, Rawlsians would not have to resort to comprehensive doctrines in order to reach a conclusion on whether and which foetuses ought to be saved, for they could resort to procedural solutions.

At this point, it might be objected that Williams's conclusion threatens Rawlsian contractualism yet again: for if Rawlsian contractualism is consistent with a number of positions on what should happen with removed foetuses, and if some of these positions suggest that these late-term removed foetuses should be destructed, it remains possible that

Rawlsian contractualism could justify that destruction, as long as it is mandated by a democratic decision.

I suggest, however, that there is an important publicly justifiable reason that does limit the range of views that would be consistent with Rawlsian contractualism. This reason is the interest in developing one's moral powers, which was fully explored in the previous chapter. As we have seen, all beings who have the potential to become persons in the political sense have this interest. And given that political liberals cannot appeal to a comprehensive doctrine such as a metaphysical conception of personhood to differentiate between children and foetuses, it follows that if children have this interest, then so do foetuses. This interest gives us a very strong *pro tanto* reason in favour of satisfying it, which can only be outweighed by current persons' very basic rights and interests. We have seen that the case of abortion is indeed a case in which this interest is outweighed by women's basic right to bodily integrity. This right serves to justify *removing* a foetus from a woman's body. Once a foetus has been removed, however, it seems unlikely that its interest in developing its two moral powers could be outweighed by current persons' basic rights and interests, for there are very few conceivable cases in which satisfying this interest would require an infringement of current persons' basic rights and interests.

To be sure, there are other important public reasons that are at play in public debates on what should happen to unwanted foetuses. For instance, relevant reasons would include population issues, the medical costs of providing treatment to premature infants, the number of parents that would adopt children, the relative social costs of increasing the number of children's homes, the social effects of not saving viable foetuses, and the value and form the family has in each society.

Although these are important considerations, given that they do not threaten current persons' basic rights, it seems unlikely that they could outweigh removed foetuses' interest in developing their two moral powers, unless they are taken to be in their most extreme form. To see why, consider, for instance, the claim that whether an ill child should be given medical assistance depends on whether there is a sufficiently low population or whether the social effects of not saving it would not be disastrous. This claim seems so counterintuitive and morally reprehensible precisely because a child's basic interest in developing the two moral powers could not be outweighed by these kinds of social considerations, which do not affect any other person's rights. Yet the same argument that leads us to conclude that there is a duty to offer medical assistance to that child also leads us to conclude that a viable (hence late-term) foetus should be given medical assistance for the same reasons. Thus, we see that Rawlsian contractualism would not be objectionably inconclusive on the issue of what should happen to removed foetuses: that is, contrary to the worry that Williams's argument opens up, the theory would not generally be consistent with the counterintuitive claim that removed foetuses should not be given medical assistance.

There is, however, one qualification to my suggestion that removed foetuses should be given medical assistance. This qualification is that there is an overriding duty to provide such medical assistance only when it can actually enable the foetus to develop the two moral powers to the requisite degree. This qualification has two intuitively plausible implications. First, it implies, in line with commonly held views, that there *is* a difference between early-term and late-term foetuses. This distinction is that early-term foetuses that have been removed are less likely to be viable and, even more importantly, are less likely

to have the potential to develop the two moral powers. By contrast, late-term foetuses have reached a stage of development that makes it more likely for them to develop the two moral powers, especially if they receive appropriate medical treatment. Thus, the distinction that even pro-choice advocates often draw between early-term and late-term foetuses would be drawn in a natural way. Second, and relatedly, the qualification implies that when an abortion (that is, the removal of a foetus) is likely to lead to that foetus's inability to develop the two moral powers (see McMahan 2006), there is no overriding duty to offer medical assistance to the foetus. In such cases, the public reasons explored above, such as the number of prospective adoptive parents, would be relevant.

We can therefore conclude that the extreme permissibility objection to Rawlsian contractualism misfires. For, even if it does have radically permissive implications for abortion, these only ground the permissibility of removing foetuses – not of ending their development. In fact, we have seen that the interest that foetuses have in developing the two moral powers ensures that Rawlsian contractualism does not have the counter-intuitive implications that Williams suggests; in line with commonly held intuitions, the removal of late-term foetuses would typically be consistent with the continuation of their development.

Still, Williams might press premise 5, which holds that permitting late-term abortions is deeply problematic. Given that I have disputed the implications that Williams ascribes to Ralwsian contractualism, it might seem that these objections have already been implicitly addressed. It remains likely, however, that even if one accepts my argument about the distinction between removing foetuses and ending their development, the objections might

still be raised against Rawls's theory.[40] That is, it might still be argued that permitting the *removal* of a late-term foetus is problematic.

The first way in which Williams presses premise 5 – termed the incongruity objection – is that, by permitting late-stage abortions, Rawlsian contractualism fails to satisfy a central desideratum: that its implications be 'within the leeway allowed by citizens' reasonable comprehensive doctrines' (Rawls 2005, 246; Williams 2015, 50). In other words, Williams argues that the fact that most citizens' comprehensive doctrines would strictly oppose public reason's conclusions on abortion should make us sceptical about the plausibility of the theory. Political liberals would surely not want to endorse a theory if they knew that the vast majority of citizens would disagree with its implications for issues of basic justice or constitutional essentials. Given that all pro-life and most pro-choice citizens' comprehensive doctrines lead them oppose late-term abortions, Williams concludes that public reason is vulnerable to the incongruity objection.

However, many citizens oppose late-term abortions mainly because they assume that the removal of a foetus also implies the end of its existence. But as we have seen, even if Williams's argument is correct, Rawlsian contractualism does not give rise to the latter implication. For women's rights only justify the foetus's removal, not the ending of its existence. One might press the objection, however, against my argument that removal might be justified. For, someone might claim, given that some foetuses will be removed before they are viable, that their removal would be opposed by pro-life comprehensive doctrines that view foetuses as rights-holders.

[40] Indeed, in private communication, Williams holds that the fifth premise applies to my proposed solution as well.

It may be doubted, however, whether a view that is in favour of violating women's basic rights is reasonable. If reasonableness requires understanding and complying with the requirements of justice and if respecting women's basic rights is part of the requirements of justice, a view that is in favour of violating those rights is unreasonable. Of course, this is not to say that the comprehensive doctrines that include this view, such as certain religious doctrines, are to be viewed as wholly unreasonable, insofar as they are also compatible with a reasonable view on women's rights. So, we can conclude that, although the incongruity objection could indeed be raised against my argument about the implications of Rawlsian contractualism for abortion, it fails to undermine the theory.

Williams gives two further reasons for premise 5, which holds that permitting late-term abortions (i.e. removals) is deeply problematic. First, he raises what he calls the 'anti-democratic objection', which states that Rawlsian contractualism is too decisive on what counts as a reason before citizens get to express their views. In this case, the problem is posed by the fact that 'political conceptions of justice exhibit absolutely no diversity on the issue of prenatal moral status' (Williams 2015, 51). Second, he mounts the 'integrity objection', which states that, by being required to exclude their views on prenatal moral status from public discourse, citizens 'alienate themselves from their deepest moral convictions, thereby compromising their moral integrity' (Williams 2015, 51).

However, as far as public reason's implications for abortion are concerned, it seems that once again the two objections lose their strength once we realise that the theory does not require the death of aborted foetuses. Again, the reason why many citizens would object to Rawlsian contractualism given Williams's argument about its implications for abortion is

presumably that they believe that sufficiently developed foetuses must be protected in some way. Indeed, the more a foetus resembles an infant, the more objectionable most people think that abortion is. Once those who object to public reason on the grounds that even developed foetuses will cease to exist realise that public reason does not, in fact, require the foetus's death, their opposition will surely weaken. This seems all the more likely given that medical and technological advancements can often enable removed foetuses to go on to live full, healthy lives. The anti-democratic objection thus loses its force, as those who complain about public reason's lack of diversity on prenatal moral status because they believe that late-term foetuses ought to be protected will be reassured that *some* such protection may be granted. The integrity objection is similarly weakened, since, presumably, the 'deepest moral convictions' that lead some citizens to oppose late-term abortion identify reasons not to *kill* the foetus (Williams 2015, 51). If, however, my argument is sound, then such reasons would not give citizens an overriding reason to object to public reason's implications for abortion.

Williams might claim that his objections would still apply to my argument, since some citizens would think that the law on abortion ought to be formulated according to arguments about prenatal moral status and since, as we have seen, many would oppose removal if it was likely to cause prenatal harm. In particular, those who believe that the foetus is a person with rights equal to those of the mother, either throughout or at the end of the pregnancy, might mount both the anti-democratic and the integrity objections. First, they might argue that the law must protect individuals' rights; any arguments on whether the foetus is a person with rights must therefore have a central place in public discourse and constitutional deliberations on the matter. Since Rawlsian contractualism declares these arguments inappropriate in the relevant context, it can be said to be anti-democratic. Second, given

that arguments about the kind of entity a foetus is are central to many individuals' 'deepest moral convictions', it may be argued that Rawlsian contractualism compromises these citizens' integrity by requiring them not to advance these arguments in the political arena.

It should be stressed here that political liberalism does not *force* citizens to only express public reasons. The non-enforceable duty of civility is a moral one, in the sense that reasonable citizens, who are committed to the idea of society as a fair scheme of cooperation, must refrain from appealing to reasons that are not justifiable to other reasonable and rational persons. In fact, Rawls's later discussion of that duty grants that citizens may appeal to their comprehensive doctrines, 'provided that in due course public reasons, given by a reasonable political conception, are presented sufficient to support whatever the comprehensive doctrines are introduced to support' (Rawls 2005, li-lii). Thus, all citizens have the freedom to defend arguments that appeal to non-public reasons, even if that means that they would be violating the moral duty of civility.

Even so, someone might raise the anti-democratic and integrity objections by claiming that the duty of civility itself represents an anti-democratic and alienating imposition of moral rules on citizens' lives. However, the fact that some citizens believe that appropriate justifications of the law on abortion should appeal to prenatal moral status does not suffice to give the two objections much force. This is because public reason liberalism does not and should not include a commitment to respecting all views on which reasons ought to be recognised as justifiable. That is, the commitment to impartiality in the face of reasonable disagreement in citizens' reasonable comprehensive doctrines does not apply to public reason liberalism itself. Thus, raising the two objections as if they are related to the implications of public reason for abortion is question-begging.

By way of illustration, notice that the same objections could be raised regarding the permissibility of grounding a pro-life law on a clearly religious basis. A significant number of citizens, for example, believe that foetuses should never be aborted because doing so is contrary to God's commands. *Any* secular theory that excludes such bases as legitimate justifications would be equally open to the anti-democratic and the integrity objections. For instance, since liberal states typically distinguish the validity of metaphysical from religious claims, religious citizens could raise the same objections against pro-choice laws that argue that non-viable, non-sentient foetuses are not persons. Similarly, secular, liberal states would not accept substituting evolutionary theory with creationist accounts in the teaching of science; the fact that these requests are excluded does not make those states anti-democratic. Their exclusion would only have that effect if we had good reasons for including them.[41] Generally, if the anti-democratic and integrity objections apply to public reason's exclusion of certain categories of justification (and not merely to specific justifications), then comprehensive doctrines are viewed as possible grounds for a law on issues of basic justice. However, this claim rejects public reason in principle, not on the basis of its implications for abortion. If public reason theorists are right in claiming that equal respect requires that the law on abortion not appeal to comprehensive doctrines, then those who disagree with that view are not being wronged by it, even if they believe that they are. If there are good and decisive reasons to appeal only to public reasons, the mere fact that those who will have to change their behaviour will object to doing so does not count as a reason against public reason. Given my assumption that the main, abstract

[41] It should be noted here that these reasons are not good in an objective, perfectionist sense. Rather, in the context of public reason, when referring to issues of basic justice or constitutional essentials, only public reasons can be 'good' reasons. This statement should not, therefore, be conflated with similarly phrased perfectionist statements about reasons.

commitments of public reason are plausible, the anti-democratic and integrity objections cannot be raised here as a critique of the theory's implications for abortion.

3.4 The Extreme Non-Permissibility Argument

David Shaw claims that Rawls's permissive views on abortion in *Political Liberalism* are not consistent with his use of the Original Position.[42] Following Williamson Evers, Shaw argues that, in the Original Position, it would be irrational for us to decide to permit abortion, as 'it seems unlikely that we would decree our own execution' (Evers 1978, 111). Given that the Veil of Ignorance prevents us from knowing our exact natural and social characteristics, it is argued that we would ban abortion out of fear that we end up being foetuses (Shaw 2011, 96). According to Evers and Shaw, the response that only adults are represented in the Original Position is not available here, as Rawls clearly states that the representatives must also consider and protect those who merely have the potential to develop the reasonable and rational capacities (Evers 1978, 111; Shaw 2011, 97; Rawls 1971, 509). Although Rawls makes this argument in order to protect children and future generations, the argument implies that foetuses ought to be protected as well, given that they similarly have the potential to develop the two moral powers (Shaw 2011, 97). Indeed, critics could further argue that the same argument applies to zygotes and gametes; if the extreme non-permissibility argument is sound, Rawlsian contractualism would actually require us to maximise reproduction, protect and develop sperm and eggs as much as possible, and ban contraception. Thus, the argument concludes that Rawlsian

[42] Although the objection rightly highlights certain ambiguities in Rawls's writings regarding the duties that persons, in the political sense, have towards non-persons, it is worth noting that Shaw's two arguments in *Justice and the Fetus: Rawls, Children, and Abortion* are not compatible. On the one hand, the indeterminacy objection claims that political liberals cannot decide on abortion using solely political values, and that they therefore implicitly rely on comprehensive doctrines; on the other hand, according to his extreme non-permissibility argument, Rawls's arguments and methodology imply that abortion is not permissible. If the latter is true, however, political liberals *can* decide on abortion using solely political values, but, contrary to what they usually argue, their conclusion has to be against abortion.

contractualism could never have pro-choice implications for abortion, much less the extreme implications that Williams identifies; rather, the theory is viewed as having extremely *restrictive* implications, given that abortion would be against the rational self-interest of the parties in the Original Position.

Before proceeding to respond to this objection, we should notice that, as it stands, the objection mischaracterises the source of the problem: while Rawls indeed claims that the parties in the Original Position ought to consider future generations' interests, we have seen in the first chapter that this is justified by appeal to reasonableness, not rationality.[43] That is, the parties do not consider future generations because they might end up as members of those generations; clearly, they do not themselves represent future generations and their decision is not driven by their self-interest. Rather, the idea of reasonableness includes the idea of temporal impartiality – a concern for all persons that will exist and because the Original Position is designed to model reasonableness, certain external constraints must be imposed (such as the Veil of Ignorance). As we have seen, one of these constraints is that the parties ought to consider future persons' interest in developing the two moral powers. Therefore, the correct interpretation of the extreme non-permissibility argument is that abortion is inconsistent with the idea of reasonableness and the associated commitment to temporal impartiality – and not that it is inconsistent with the parties' desire to pursue their self-interest in the Original Position.

[43] Rawls actually argues that 'the correct principle is that which the members of any generation would adopt as the one their generation is to follow and as the principle they would want preceding generations to have followed (and later generations to follow), no matter how far back (or forward in time)' (Rawls 2005, 274). Clearly then, this may not be decided by mutually disinterested, rational parties. Rather, it is a requirement of the reasonable and ought to be somehow incorporated into the methodology of the Original Position, which is formulated according to those requirements.

Indeed, as we saw in the first section of this chapter, there seems to be tension between my argument that all future persons have an interest in developing the two moral powers and the claim that abortion is morally permissible. However, there are two reasons why this tension is not sufficient to ground the extreme non-permissibility argument. First, as I have argued, removing a foetus does not necessarily imply that its development will end. A foetus's interest in developing the two moral powers can therefore often be satisfied even if it has been removed. This means that a permissive stance on abortion does not necessarily conflict with the view that foetuses have an interest in developing the two moral powers which often gives rise to duties to fulfil that interest. Yet it is true that there are some cases in which there is indeed a conflict. These are cases in which the foetus's life cannot be sustained once it has been removed. This brings us to the second reason why the extreme non-permissibility argument does not hold. As we briefly saw in the first section, although those who currently count as persons have a duty to enable future persons to develop their two moral powers, this duty does not override all other duties and rights; this is because the corresponding interest that future persons have in developing the two moral powers is typically outweighed by the very basic rights of those who currently count as persons. In the case of abortion, preventing women from removing foetuses from their bodies when they wish to do so violates their basic right to bodily integrity. Given that basic rights outweigh future persons' interests, we can conclude that removing foetuses is permissible even if the foetus's development will end as a result; that is, abortion is permissible even if it is inconsistent with the fulfilment of the duty that current persons would otherwise have to enable the development of future persons' moral powers.

3.5 Conclusion

This chapter has considered the most prominent critiques of Rawlsian contractualism's implications for abortion. First, according to the incompleteness objection, public reason is either indeterminate or inconclusive. The charge of indeterminacy states that public reasoners *cannot* say anything about abortion because they exclude the issue of metaphysical personhood from the considerations that they can appeal to. The inconclusiveness objection, on the other hand, claims that public reason is consistent with too many different positions on abortion, and cannot, therefore, provide us with a clear conclusion on what we ought to permit. Second, according to the extreme permissibility argument, public reason implies that abortion should be permitted even in late-term pregnancies; this is because public reason cannot say anything on prenatal moral status and can only appeal to women's rights. Third, the extreme non-permissibility argument states that, on the contrary, public reason is not consistent with any pro-choice claims. This is because the parties in the Original Position would never permit abortion, given that they might end up being aborted foetuses.

I have rebutted these objections in a number of ways. A key premise of all of my arguments has been that abortion need not always imply the foetus's death. Since women's rights, as specified on a political liberal view, only justify the *removal* of a foetus, Rawlsian contractualism's implications on the question of abortion do not straightforwardly imply anything about whether foetuses continue to live. This is particularly the case as medical technologies progress. It follows that, even if some of the objections examined in this chapter are correct, it does not follow that Rawlsian contractualism is as objectionable as it is envisaged. The only objection that might still remain is that of inconclusiveness; that, however, does not seem to be particularly problematic for this (or any) theory.

Chapter 4

Unreasonable Persons

4.1 Introduction

As we have seen, Rawlsian contractualism faces an important question of exclusion. Briefly, the problem is that because the theory's commitment to public reason prevents it from relying on comprehensive doctrines, the theory cannot appeal to metaphysical conceptions of personhood or moral status to determine who counts as a person to whom states and citizens have enforceable duties. In order to resolve this issue, Rawlsian contractualism must resort to a *political* conception of the person, which is defined with reference to the political idea of social cooperation. More specifically, in order to count as a person one must be capable of participating in social cooperation – which means that one must be sufficiently reasonable and rational. Reasonableness is meant to capture an epistemic power - the capacity to understand the requirements of justice – and a motivational capacity, which is the willingness to comply with these requirements. Clearly, reasonableness is a necessary aspect of the capacity to participate in social cooperation because it ensures that one can *cooperate* with others *under fair terms*. And being rational means that one possesses the capacity to form, revise, and pursue a conception of the good. The reason why rationality is necessary for one's participation in social cooperation is that it gives individuals something to cooperate *for*; given that society is viewed as social cooperation with the purpose of *mutual advantage*, all individuals should be motivated to participate. But the corollary of this solution to the problem of who counts as a person is

that a number of groups who do not fit the political account of personhood are entirely excluded from the theory.

In this chapter, I explore the problem of exclusion as it applies to unreasonable persons. In particular, I argue that the definition of political personhood, combined with the significance of social cooperation in Rawlsian contractualism, implies that unreasonable persons are excluded from the constituency to whom justification for the exercise of political power is owed and from those to whom the principles of justice apply. This argument is developed in the first section, which explores the ways in which unreasonable persons are excluded from Rawls's theory of justice and legitimacy.

In the second section, I examine the kinds of problems the exclusion of unreasonable persons causes for Rawlsian contractualism. In particular, I argue that this exclusion makes the theory vulnerable to the objections that it is internally inconsistent, indeterminate, and that it has counterintuitive implications.

I then turn, in the third section, to suggest a way in which unreasonable persons can be included in Rawlsian contractualism without abandoning the commitments of public reason, including the political conception of the person. I argue that there are three main reasons why political liberal states have a duty to enable unreasonable persons to develop their reasonableness. These are that these persons have an interest in developing their two moral powers, that the existence of unreasonableness threatens citizens' basic rights, and that unreasonableness further threatens society's stability. As with other groups, such as children and future generations, this way of including unreasonable persons suggests that they are included in Rawlsian contractualism under different terms from those under which

reasonable persons are included. For instance, their inclusion does not imply that they have, here and now, full participation rights in the justificatory constituency. Their interests are represented in that constituency, however, in a way that ensures their future participation.

In the fourth section, I examine the implications of this way of including unreasonable persons in Rawlsian contractualism. I claim that my argument implies that mandatory education, therapy, and even neurointerventions – that is, interventions that act directly on one's brain, such as Deep Brain Stimulation or anti-libidinal agents – are often justified. I devote special attention to the latter kind of interventions, which have not been much explored by others and are moreover both particularly controversial.

Finally, I consider three pressing objections. According to the first objection, if states are responsible for unreasonableness, given that it is often caused by injustice, there seems to be something problematic about suggesting that states ought to treat unreasonable persons in ways (such as imposing non-consensual paternalistic interventions on them) that would never be permissible with regards to reasonable persons. According to the second objection, my argument has the very implausible implication that some humans are less than equal. That is, the objection claims that, according to my argument, the moral worth of some adults, such as unreasonable persons, is lesser than that of all those who count as free and equal citizens. Finally, the last objection raises the claim that if we accept one of the policies that I defend as an implication of my argument – non-consensual neurointerventions – we must also accept a clearly counterintuitive implication, namely, that indoctrination is similarly justified. I argue in response that, although all three objections highlight important concerns, they do not undermine my argument.

4.2 The Exclusion of Unreasonable Persons

4.2.1 Exclusion from the Justificatory Constituency

The ways in which unreasonable persons are excluded in Rawlsian contractualism will be familiar from the preceding chapters on children and foetuses. Given the commitments of Rawlsian contractualism, as the theory stands, unreasonable citizens seem to be excluded in two ways. First, they are excluded from the constituency of those to whom justification for the exercise of political power is owed. To see why unreasonable citizens are so excluded, recall the liberal principle of legitimacy:

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

In order to count as a free and equal citizen, however, one must be sufficiently reasonable and rational. One must be reasonable, in that one should have the capacity to understand the requirements of justice and the willingness to comply with these requirements, and one must be rational, in that one should have the capacity to form, revise, and pursue a conception of the good. As we have seen, the two moral powers are required in order to characterise a person as free for two reasons. First, reasonableness and rationality are necessary conditions for one to be able to view oneself as an independent and 'self-authenticating source of valid claims' (Rawls 2005, 32). Recall, for example, a case we saw in Chapter 1, of someone who is rational but unreasonable. Despite her having the capacity to form, revise and pursue her conception of the good, the self-authenticating claims that would arise from that unreasonable conception would not be *valid*, given that they would not be consistent with justice. Similarly, a person who is reasonable but irrational would lack the capacity to be a *self-authenticating source* of a claim regarding

the way in which she should lead her life. That is, even if she did endorse such a claim, it would not be appropriately viewed as her own. A person who lacks rationality so construed would not be the *author* of the claims and doctrines that guide her life. The second reason why the two moral powers are necessary conditions for freedom is that these capacities enable citizens to view themselves as responsible for the ends and doctrines that they form and pursue (Rawls 2005, 33).[44] Notice, for example, that we would not hold an infant morally responsible for any ends that it formed or pursued because infants normally lack the capacities that are necessary for moral responsibility. If being morally responsible for our choices is part of what it means to be a free person who can make *free* choices, we see that the possession of the two moral powers is a necessary condition for counting as a *free* person. Indeed, this is why, for example, infants and non-human animals are not viewed as free persons; because they lack the capacities that are necessary in order to be viewed as independent and responsible self-authenticating sources of valid claims.

Apart from being a necessary condition for freedom, we have seen that the possession of the two moral powers is further a necessary condition for equality (Rawls 1971, 441—449; Rawls 2005, 19; Rawls 2001, 20). This is not entirely specific to Rawls's theory. When addressing the question of why certain individuals are equal and why others – often including foetuses, children, and non-human animals – are not, it is often held that one or another cognitive capacity forms the basis of individuals' equality. The reason why the relevant capacities are *reasonableness* and *rationality* in Rawlsian contractualism is that these capacities enable individuals to participate in social cooperation for mutual advantage. Reasonableness enables them to *cooperate* with others under *fair* terms, while

[44] It should be clear that this conception of freedom does not imply that individuals are *metaphysically* prior to or responsible for their ends.

rationality enables them to pursue their rational advantage and thus participate in cooperation that aims at *mutual advantage*. It is for this reason that reasonableness and rationality form the basis of equality. Thus, we see that if only reasonable and rational persons count as free and equal citizens, and if only free and equal citizens are included in the constituency to whom justification for the exercise of political power is owed, it follows that unreasonable persons are excluded from that constituency (Friedman 2000, 16; Quong 2011, 298).

However, there might seem to be something objectionable about the exclusion of unreasonable persons, since that exclusion effectively states that 'a state that is endorsed only by its reasonable citizens is thereby sufficiently entitled to exercise its coercive power over *unreasonable* citizens *without* their consent' (Friedman 2000, 16). I return to the question of what is problematic about the exclusion of unreasonable persons in the next section; at this point, however, the pertinent issue is that we can conclude that unreasonable persons are indeed excluded from the constituency of legitimacy – that is, the constituency to whom justification for the exercise of political power is owed.

4.2.2 Exclusion from the Constituency of Justice

Unreasonable persons seem to be excluded in an even more important way: they seem to be excluded from the constituency of justice, which is to say that they are not to be viewed as agents of justice and that the principles of justice do not apply to them.[45] If this is true, then unreasonable persons are not protected by the principles of justice. Taking the

[45] As we saw in the chapter on children, with the exception of a few principles, such as the natural duty of justice and the principle of fairness, the principles of justice apply to the basic structure of society. By claiming that the principles do not apply to certain individuals such as unreasonable persons, I mean that the rights protected by the two principles of justice are not assigned to these individuals *and* that the principles that do apply to individuals do not apply to these specific persons.

Rawlsian theory of justice as the relevant theory, this kind of exclusion means that the basic liberal rights and liberties of unreasonable persons are not to be protected and that the state is not under a duty to promote unreasonable persons' opportunities nor to protect their interests in accordance with the difference principle.

The reason for this exclusion is fairly straightforward. Any theory of justice that is consistent with the commitments of Rawlsian contractualism aims at the fair distribution of the benefits and burdens of social cooperation among those who have the capacity to participate in social cooperation. For instance, the theory is not concerned with the distribution of the benefits and burdens of humans' social cooperation among *non-human animals*, precisely because they cannot participate in social cooperation.[46] It is for this reason that the principles of justice only apply to humans and that, consequently, states and citizens do not have an enforceable duty to respect animals' basic rights and liberties as the liberty principle instructs them to do with regards to humans. It is for this reason that the difference principle does not protect animals (Rawls 1971, 441). Since the justification for the exclusion of non-human animals is that they lack the two moral powers which are necessary for one's participation in social cooperation, and since unreasonable persons similarly lack at least one of the two moral powers, it seems that if animals are excluded from the constituency of justice, then so are unreasonable persons. Indeed, Rawls seems to admit this point, without acknowledging the implications for unreasonable persons, when he states that 'it is precisely the moral persons who are entitled to equal justice…We use the characterization of the persons in the Original Position to single out the kind of beings to whom the principles chosen apply' (Rawls 1971, 442). Thus, not only are unreasonable

[46] Admittedly, some argue for a wider interpretation of contributions to social cooperation, which would allow non-human animals to be included in those who *can* contribute. I discuss this view in the final chapter.

persons not included in the constituency to whom justification for the exercise of political power is owed, but they are also excluded from the group that is protected by the principles of justice.

Jonathan Quong has argued against this reading of Rawlsian contractualism, claiming that Rawls's principles of justice, which include the liberty principle, protect unreasonable persons too. In his words,

> it is fairly straightforward: if I accept (a) that all citizens are free and equal, (b) that the original position models this premise correctly, and (c) that the principles of justice derived from the original position are correct, then I must accept (d) that those principles apply to all citizens. The mere fact that someone denies the truth of (a) is not a reason not to treat them as a citizen, and thus not a sufficient reason to deny their basic rights or other liberties (Quong 2011, 293).

However, we may doubt (a), given that the Rawlsian account of freedom and equality implies that not *all* citizens or individuals are free and equal. This is because, as we have seen, a necessary condition for persons' freedom is that they have sufficiently developed the two moral powers – the reasonable and the rational – and the powers of reason, that is, the powers of judgment, thought, and inference (Rawls 2005, 19). Similarly, we have seen that 'their having the moral powers to the requisite minimum degree to be fully cooperating members of society makes persons equal' (Rawls 2005, 19). Possessing the two moral powers is taken to be 'the *necessary* and sufficient condition for being counted a full and equal member of society' (Rawls 2005, 310, my emphasis). It follows that those who do not possess the two moral powers to a sufficient degree do not pass the equality test.

In fact, we generally accept that children are not automatically granted the rights and liberties that are typically accorded to free and equal citizens, precisely because they lack

the capacities that ground political personhood. The reason why we treat children differently from adults and do not ascribe to them exactly the same rights and duties that adults have is that we view age as a proxy for the extent to which one possesses certain cognitive and moral capacities; taking Rawlsian contractualism as the relevant framework, these capacities would be reasonableness and rationality. If human development were different and we knew that humans typically develop the two powers to the requisite degree at an earlier age, we would not be justified in treating individuals of that age as less than equal citizens.

Similarly, then, unreasonable persons who fall below the reasonableness threshold may, consistently with political liberalism's premises, turn out to lack at least some of the rights and liberties that citizenship guarantees. In the absence of further argument, there seems to be no good reason to assume that persons who lack the reasonable moral power can have the status of free and equal citizens who enjoy *all* the rights and liberties associated with that status. In these cases, therefore, it is not clear why (a) would be taken to be true; and if (a) is not true, then it is not clear why political liberals should accept Quong's conclusion (d) – that the principles of justice, including the protection of basic liberties, apply to all citizens, including unreasonable criminal offenders and children. So, for all that Rawlsian contractualism says, unreasonable persons may be excluded from the constituency to whom the principles of justice apply.

4.3 Why Is Exclusion a Problem?

I have suggested in previous chapters that the exclusion of different groups from the constituencies of justice and legitimacy gives rise to serious problems that threaten Rawlsian contractualism. And so it proves in the case of the exclusion of unreasonable

persons as well. The first problem that the question of unreasonable persons' exclusion raises is that it makes Rawlsian contractualism internally inconsistent and independently objectionable. This is because many unreasonable persons have the potential to develop their reasonableness, but this potential has not been realised for a number of reasons, such as, for example, the upbringing they happened to have had. As a result, the theory is effectively biased against individuals who have happened to face bad social luck. But Rawlsian contractualism is otherwise opposed to unfair bias in favour of those who have had good social luck. For instance, the reason why having unequal opportunities as a result of bad parenting is bad is that having better life prospects as a result of social luck is entirely undeserved and, hence, unfair. As a result, not developing the two moral powers due to social luck when one otherwise had the potential to develop them is a clear instance of unfairness.[47] And a theory of justice that would favour those who have better social luck in this sense would be a theory that reveals unfair bias. Given that Rawlsian contractualism is otherwise opposed to such unfair bias, the exclusion of unreasonable persons makes the theory vulnerable to the charge of internal inconsistency. And to the extent that favouring those who have happened to develop their potential is indeed an instance of *unfair* bias, Rawlsian contractualism is further vulnerable to the external criticism that it is an implausible theory of justice and legitimacy.

The second problem that Rawlsian contractualism faces as a result of excluding unreasonable persons from the constituencies of justice and legitimacy is the threat of indeterminacy. The worry here is that, because unreasonable persons are excluded,

[47] The reason why my claim is restricted to those who have the *potential* to develop the two moral powers but do not is that those who lack this potential are not included in social cooperation as current or *future* persons. As a result, the concept of fairness (in the distribution of the benefits and burdens of social cooperation) cannot apply to them in the first place. I address this issue in the chapter on non-human animals.

Rawlsian contractualism is not consistent with *any* views on the question of what duties, if any, are owed to them. This is an implication of the theory's reliance on political liberalism and the related idea of public reason, which require that issues of basic justice and constitutional essentials be justified in political, rather than comprehensive, terms. According to the indeterminacy objection, the public reason restriction implies that Rawlsian contractualism cannot justify either the existence or the absence of any enforceable duties towards those who do not count as political persons without implicitly appealing to a comprehensive doctrine, such as a metaphysical view of personhood, on the validity of which there can be reasonable disagreement. As a result, it is argued, Rawlsian contractualism is *indeterminate* on issues that concern non-persons, because reaching *any* conclusion on these issues requires implicitly appealing to a comprehensive doctrine. Indeed, for example, this is why, as we have seen, Rawlsian contractualism has been accused of being indeterminate on the issue of abortion. I argue that the same worry is relevant in the case of unreasonable persons; that is, since unreasonable persons are not persons in a political sense, and if Rawlsian contractualism can *only* reach conclusions for political persons, then the theory (and, by extension, political liberal states) cannot reach any conclusions on matters that concern unreasonable persons. We thus see that the question of unreasonable persons' exclusion leaves Rawlsian contractualism vulnerable to the charge of indeterminacy.

Finally, the exclusion of unreasonable persons from the theory, particularly with regards to the constituency of justice, raises the objection that Rawlsian contractualism has extremely counterintuitive implications. Indeed, as we have seen in response to Quong, the fact that unreasonable persons do not fit the account of political personhood seems to imply that they do not possess any rights or basic interests that can be recognised within a political

liberal theory of justice and that, relatedly, states and citizens do not have any enforceable duties towards them. Clearly, this implication seems extremely counterintuitive. Given the role of reflective equilibrium in Rawlsian contractualism, the extreme counterintuitiveness of the theory's implications for the political status of the unreasonable presents the theory with a serious challenge (Rawls 2005, 246).

Consequently, if unreasonable persons are indeed excluded from the constituencies of justice and legitimacy in Rawlsian contractualism, the theory faces the problems of internal inconsistency, indeterminacy, and counterintuitive implications. In order to avoid these pressing problems, Rawlsians must find a way to respond to the question of exclusion as it applies to unreasonable persons. This is what I do in the following section.

4.4 Including Unreasonable Persons in Rawlsian Contractualism

I suggest that Rawlsians can include unreasonable persons in Rawlsian contractualism without abandoning the commitments of public reason by focusing on the duty that I defended in chapter 2, which requires political liberal states to enable unreasonable persons to develop their reasonableness to the requisite degree. There are three reasons supporting this duty, which are all consistent with the requirements of public reason. Although the interest in developing the two moral powers is typically only affirmed with regards to children, I suggest that the same interest and the duty that arises from this interest is in place with regards to unreasonable persons. This claim and its interesting but controversial implications provide us with an approach to the problem of unreasonable persons that is very different from the standard approach in the relevant literature (see Rawls 2005, 64; Quong 2011, 293; Friedman 2000).

The first reason in support of this claim is that, as we saw in the chapter on children, everyone who has the potential to develop the two moral powers has a basic interest in developing them; political liberal states, in turn, have a duty to fulfil this basic interest. As we have seen, the reason why we all have an interest in developing the two moral powers is that we have a basic interest in participating in social cooperation and the two moral powers are necessary conditions for that participation. It might seem that arguing that non-persons can have a basic interest in becoming persons that further gives rise to enforceable duties is inconsistent with the requirements of public reason. I contend, however, that this is not the case because our fundamental interest in participating in social cooperation which grounds our interest in developing the two moral powers is a *political* one, given that it is justified with reference to the *political* idea of social cooperation. Regardless of one's conception of the good, we all benefit from participating in social cooperation given that, by definition, this cooperation aims at citizens' *mutual* advantage. Simply put, in the absence of social cooperation, we would lack the institutional framework within which we can pursue our conception of the good without worrying about the violation of our basic rights and liberties. If the development of the two moral powers is necessary for the receipt of the benefits of social cooperation, all beings who have the potential to develop the two moral powers have a basic *political* interest in developing them.

In order to see why unreasonable persons have a basic interest in developing their reasonableness which gives rise to enforceable duties, we can further rely on an analogy with children. The thought here is that, much like unreasonable persons, children face a potential issue of exclusion from the constituencies of justice and legitimacy for the same reason that unreasonable persons do – that is, because they lack the two moral powers. Yet, as we have seen, the way for Rawlsian contractualism to include children in the theory is

by endorsing the claim that children have a basic interest in developing the two moral powers, which ought to be taken into consideration by those included in the constituencies of justice and legitimacy. Indeed, Rawls and other Rawlsians often appeal to this interest in passing. For instance, Rawls argues that

> the aim of the parties [in the original position] is to agree on principles of justice that enable the citizens they represent to become full persons, that is, adequately to develop and exercise fully their moral powers and to pursue the determinate conceptions of the good they come to form (Rawls 2005, 77).

Similarly, Quong implicitly appeals to our interest in developing the two moral powers when he claims that, although parents have the right to make educational decisions for their children, states ought to intervene when the education provided is inconsistent with states' duty to 'provide a fair framework within which citizens can develop and exercise their moral powers' (Quong 2011, 305). If we accept that children indeed have an interest in developing the two moral powers and that this interest gives rise to correlative duties, then we must also accept that unreasonable persons may have the same interest, which gives rise to similar duties. Thus, the analogy with children shows that unreasonable persons *can* be included in Rawlsian contractualism; although they are not included in the same way as those who are *currently* reasonable and rational, their basic interests are represented in the constituencies of justice and legitimacy.

The second reason supporting political liberal states' duty to fulfil unreasonable persons' interest in developing their reasonableness is given by states' duty to protect the basic rights and interests of those who currently count as persons in the political sense – that is, sufficiently reasonable and rational persons. Given that unreasonable persons are, by definition, either unable to understand the requirements of justice or unwilling to follow

them, the existence of unreasonableness threatens the basic rights and interests of reasonable and rational persons. Assuming that the most effective way of eradicating unreasonableness is to ensure that all unreasonable persons become reasonable, a political liberal state's duties towards its citizens give us a second reason for the claim that states have a duty to cultivate all persons' reasonableness.[48]

Finally, there is a third reason supporting the same claim that appeals to the value of stability. This justification for the duty that states have towards unreasonable persons starts with the observation that unreasonable persons may threaten society's stability. In particular, unreasonable doctrines and acts that aim to violate individuals' rights can limit citizens' trust in the state's effectiveness in protecting their basic rights and liberties. Similarly, the potential influence of unreasonable persons on others, including those who are particularly malleable, such as children and other persons who have not yet developed the two moral powers, may lead to a less stable society. Given that Rawlsian contractualism requires stability for the right reasons (that is, the *reasoned* allegiance of a state's citizens – as opposed to a mere *modus vivendi*), increasing numbers of unreasonable persons or doctrines can threaten to undermine the shared political values of freedom, equality, and fairness. In those cases, 'a normatively stable liberal democratic regime becomes impossible' (Quong 2011, 300). Indeed, some political liberals have granted that, for these reasons, unreasonable doctrines may sometimes need to be contained. For example, according to Rawls, 'that there are doctrines that reject one or more democratic freedoms

48 Of course, it might be objected that the most effective way of eradicating unreasonableness is by eradicating unreasonable persons themselves. I do not consider this possibility because it seems wildly counterintuitive for a number of reasons. Some of these are that this suggestion is not consistent with the political interest in developing the two moral powers; that there is overlapping consensus among reasonable comprehensive doctrines against this position; and that endorsing the eradication of unreasonable persons would lead to the cultivation of traits, such as apathy, that are inconsistent with the political virtues and emotions required for reasonableness.

is itself a permanent fact of life … This gives us the practical task of containing them – like war and disease – so that they do not overturn political justice' (Rawls 2005, 64). Similarly, Quong agrees that 'a justification for containment can be grounded … on the fundamental importance of normative stability in a well-ordered liberal society' (Quong 2011, 300). Although the containment argument has not been fully developed by political liberals, it seems plausible to assume once again that the most effective way of containing unreasonableness is through the cultivation of reasonableness. The argument from stability then suggests a third reason why states have a duty to enable unreasonable persons to develop their reasonableness.

These three reasons collectively suggest that, even if unreasonable persons are excluded from the constituencies of legitimacy and justice as they stand, their interests need not be excluded from Rawlsian contractualism. The theory can include unreasonable persons as it includes other groups that have the potential to develop the two moral powers, such as children; that is, by focusing on political liberal states' duty to enable all persons to develop their reasonableness and rationality. To be sure, this inclusion is different from the full inclusion of reasonable and rational persons, who are granted full participation rights in the justificatory constituency and to whom *all* the principles of justice apply as they are. It is, however, an important way of including the basic interests of unreasonable persons who would otherwise be entirely excluded from the constituencies of legitimacy and justice.

4.5 Implications

What does this argument imply for policy- and law-making? The first, more obvious implication is given by the analogy with children. We see that, although paternalistic interventions (that is, interventions that are justified by appealing to specific persons'

interest in receiving them) are viewed as presumptively wrong in Rawlsian contractualism, the theory implies that they are both often permissible and, even more strongly, actually *required* in the case of unreasonable persons. For example, if education can promote one's interests or well-being, then just as non-consensual education is permissibly imposed on children, the same justification could be given in defence of imposing it on unreasonable persons. Similarly, if therapy or other rehabilitative measures can enable unreasonable persons to develop the capacities that grant one full political status as a free and equal person, my argument would imply that political liberal states have a duty to impose these therapies on unreasonable persons.[49] In general, the policy implications of my argument depend on whether in each particular case the epistemic or motivational aspect of reasonableness must be targeted.[50] For instance, education may be the appropriate way to fortify one's lack of moral understanding, but cases of motivational deficiencies may require different interventions, such as counselling.

This line of argument may also lead us to a more controversial implication: that when less invasive measures, such as education, are not effective or cost-effective, so-called 'neurointerventions' might have the effects that paternalistic policies aim to produce (Theofilopoulou 2019). In the following discussion I will focus on this implication in particular, given that these interventions are underexplored and likely to face a number of philosophical objections that would not be raised against indirect interventions such as education and therapy.

[49] Interestingly, the same argument would serve to justify therapy and similar policies for children. The main reason why this is more relevant in the case of adults is that children have not yet developed defence mechanisms or psychiatric conditions that might be preventing them from developing the capacities that grant one full political status.
[50] As we have seen in the chapter on children, the epistemic aspect of reasonableness captures the capacity to *understand* the requirements of justice, while the motivational aspect captures the capacity and willingness to *comply* with these requirements.

The most widely discussed case of neurointerventions that are currently offered or administered to individuals that would be classified as unreasonable is that of anti-libidinal pharmacological agents administered to sex offenders. There is wide evidence that agents such as cyproterone acetate (CPA), medroxyprogesterone acetate (MPA), and gonadotrophin-releasing hormone (GnRH) agonists can reduce one's testosterone levels, thereby inducing chemical castration (Gordon and Grubin 2004; Chew et al. 2018). Similarly, it has been observed that a side-effect of selective serotonin reuptake inhibitors (SSRIs), which are typically used to treat depression and anxiety, is a decrease in patients' libido, although there is no consensus on the explanation of this side-effect (Prabhakar and Balon 2010). Although there is still disagreement on the effectiveness of these agents, there are studies that indicate that they significantly decrease the probability of reoffending.[51] These effects are reversible, as maintaining them requires constant treatment (Chew et al. 2018).

As a second example, consider pharmacological agents that target serotonin levels. Lower levels of serotonin have been associated with aggression and antisocial behaviour; for instance, many individuals with aggressive and antisocial behaviour have mutations in the gene coding for the 5-HT2B serotonin receptor, or variations in the gene coding for MAOA, a brain enzyme that is crucial for breaking down serotonin (Chew et al. 2018). Although the causal relationship between serotonin and aggressive or antisocial behaviour is not completely clear, several studies indicate that a fall in individuals' serotonin increases

[51] For example, the Oregon Depo-Provera trial studied the effects of MPA agents on three groups of sexual offenders. The first consisted of offenders who underwent treatment, the second had offenders who were prescribed MPA but did not undergo treatment, and the third contained offenders who were not eligible for prescription. The trial indicated that only 5% of those in the first group re-offended, and none of the offences were sexual, while 26—30% of the other two groups re-offended, and half of the offences were sexual (Maletzky and Field 2003).

instances of aggression (Dougherty et al. 1999; Moeller et al. 1996; Cleare and Bond 1995). Moreover, those with low serotonin cannot be deterred from aggression as effectively as others can by, say, the threat of punishment because lower serotonin levels make individuals more likely to value harming others in retaliation more than they value their own well-being (Crockett et al. 2013). More importantly, for the purposes of this discussion, there is evidence that increasing serotonin by the use of SSRIs causes a correlative fall in aggression, particularly in individuals with such tendencies and backgrounds (Chew et al. 2018). When prescribed to unreasonable persons who have already violated the requirements of justice in the past, pharmacological agents that increase serotonin levels are therefore expected to develop their reasonableness, enabling them to refrain from reoffending.

Given that such interventions can promote one's interests (indeed, we already accept that non-consensual psychiatric treatment is sometimes justified) and given that their administration does not clash with other duties that states may have towards unreasonable persons, we can conclude that states may often have a duty to administer neurointerventions to unreasonable persons.

4.6 Objections

4.6.1 The Causes of Unreasonableness

At this point someone might claim that, given certain claims about the causes of unreasonableness, the fact that unreasonable persons are included under *unequal* terms, which often require policies (such as mandatory education and neurointerventions) that are otherwise unjustified, is objectionable. The thought here is that unjust actions are usually committed due to social and economic structures that marginalise certain individuals (Barn

2016). This can happen in two ways. First, it may be that someone who is not able to cover her basic needs might commit a crime, such as theft, in order to satisfy them. Second, it seems likely that someone who grows up in unjust circumstances, such as poverty and social marginalisation, might develop both mental health problems and unreasonableness in the sense that they view social cooperation as futile and undesirable. Consequently, the objection goes, given that unreasonableness is caused by injustice, it seems unfair that unreasonable persons are not *fully* included under the same terms that reasonable persons are. In other words, it seems unfair for states to subject unreasonable persons to paternalistic treatment that is otherwise seen as presumptively wrong, considering that states themselves are often responsible for those persons' unreasonableness.

It should be noted here that this objection applies to *any* kind of criminal justice intervention towards unreasonable persons, including current practices of incarceration. There are several problems with it as applied to my argument, however. First, even if the objection challenged the political liberal argument, it would only do so for cases where criminal inclinations have indeed been caused by injustice. As such, it would not apply to cases where the causes are purely biological or cases that are due to individuals' behaviour, such as an abusive household that is not tied to structural injustice in any way.

Second, an unjust state would be viewed as illegitimate from the standpoint of political liberals to begin with, given that it could not be justified to reasonable and rational citizens. Therefore, the objection unfairly assumes that a Rawlsian proponent of these interventions somehow legitimises the non-ideal injustices that cause unreasonableness. Clearly, requiring individuals to be reasonable in the sense of respecting the requirements of justice does not imply that we should require them to respect unjust laws and structures. In order

to examine the requirements of justice in these cases, we would have to explore the question of what actions are permissible to citizens when they live in non-ideal conditions, such as under the authority of an unjust state.

Third, this objection highlights the problems bound up with ordinary notions of responsibility in non-ideal circumstances. As such, it does not seem to challenge rehabilitative accounts of how states should treat unreasonable persons in the way that it does, say, retributivist ones, which attempt to show that those who violate the requirements of justice *deserve* to be punished. Therefore, interventions aimed at cultivating reasonableness may be justified even in cases where unreasonableness has been caused by injustice, since the goal is not to *punish* anyone, but to discharge a state's duties towards both unreasonable and reasonable persons. To see why, suppose, as it happens, that a state permits or even causes extreme poverty for the majority of its citizens, leaving the top one percent to enjoy far more resources than justice permits. Suppose that these citizens, due to their unfortunate and unjust circumstances, develop various mental and physical illnesses. In this case, the state clearly has a duty to offer mental and physical treatment, which may include the prescription of certain drugs. Similarly, states have the duty to offer these citizens a standard of education that would cultivate their reasonableness and rationality, in order to give them the capacity to pursue their conception of the good within the constraints set by the requirements of justice. By extension, therefore, if mandatory paternalistic policies such as education or neurointerventions can indeed increase an unreasonable person's reasonableness, states have a duty to employ these interventions, regardless of the origins of that unreasonableness. And this is not simply in virtue of the state's duty to protect other citizens or to promote stability; rather, a crucial justification of

such mandatory paternalistic policies is that they promote individuals' interest in developing their reasonableness.

This does not, in any way, imply that these states are suddenly absolved of their responsibility to strive for justice. On the contrary, justice *requires* that states offer all the means that can endow a citizen with the capacities associated with moral personality, including all primary goods. Thus, we see that requiring unreasonable persons to undergo interventions that will enable them to develop their reasonableness is consistent with the claim that states are morally responsible for the social conditions that often cause unreasonableness and that providing these interventions does not imply that states are suddenly absolved from their responsibility for creating and maintaining just social conditions.

4.6.2 The Inequality Objection

Now it might be objected that there is something deeply problematic about the claim that some individuals do not possess full political status as free and equal persons. There is something inegalitarian about this argument that seems to be in tension with the basic tenets of liberal democracies.

This objection, however, seems to conflate equal value and equal authority (see Kirby 2018). When one argues that some individuals, such as children, individuals with severe cognitive disabilities, or unreasonable persons, do not possess the capacities that confer full moral and political status, one does not necessarily have to imply anything about these individuals' value or moral worth. All that is needed is the claim that these individuals' political status does not grant them equal *authority*; that is, they are not equal sources of

self-authenticating claims.[52] It is precisely for that reason that paternalistically justified interventions are permissible in these cases; for the desires of a child, for example, do not carry the authority that the rational and reasonable desires of a free and equal adult do. Indeed, this is why Rawls himself grants that in the Original Position the parties 'will want to insure themselves against the possibility that their powers are undeveloped and they cannot rationally advance their interests, as in the case of children' by consenting to paternalistic intervention (Rawls 1971, 218).

Yet this does not imply that these individuals do not have equal *value*. For instance, although children do not possess the capacities that confer equal *authority*, their interests and rights matter as much as any other person's. In fact, it is precisely because their interests carry great moral weight that paternalistic policies that aim to promote these interests, such as mandatory education, are morally *required*. The same argument applies to the case of unreasonable persons. Claiming that their moral and political status is not inconsistent with paternalism does not imply anything about their moral worth; rather, the only implications concern their authority.

4.6.3 The Challenge of Indoctrination[53]

At this point, it might be argued that if my argument implies that it is permissible for a state to impose non-consensual neurointerventions, then it also, implausibly, implies that

[52] See, for instance, Kirby, who states that the liberal conception of equal authority has two clauses: 'First, it has a negative clause: no individual has fundamental authority over any other individual. No one is anyone else's master, no one is anyone else's subject. Secondly, it has a positive clause: each individual has fundamental authority over herself. Each individual is both her own master and her own subject. It follows that the 'liberal conception' of Equal Authority, also implies a parallel concept of 'basic freedom' also composed of two clauses, that is, first, a freedom from the fundamental authority of anyone else; and, secondly, a freedom to exercise fundamental authority over oneself' (Kirby 2018, 304).

[53] This section reproduces material from Theofilopoulou 2019, 163—165.

nonconsensual indoctrination is permissible, since this could also increase one's moral powers. If my solution to the question of exclusion has *such* counterintuitive implications, it might be argued, then we are led back to the problem of counterintuitive implications that we were meant to be solving.

Seeking to resist the permissibility of indoctrination while allowing that the purpose of criminal justice is the fortification of one's reasonableness, Jeffrey Howard argues that we need to distinguish between 'approaches that empower offenders as moral agents – that fortify their moral powers – and those that bypass them' (Howard 2017, 66). Drawing on the example of Alex in Anthony Burgess's *A Clockwork Orange*, he suggests that

> the problem with the conditioning Alex receives is not that it deprives him of his freedom of what to think – he retains the reflective capacity to affirm convictions – but that it fails to attend to the actual root of the problem: Alex's attitudes toward his fellow human beings. He is moved to refrain from violating others' rights simply because he is averse to feeling ill – not because he has grasped, and effectively been moved by, an appreciation of others' value (Howard 2017, 66).

However, there are two distinct reasons why endorsing this claim about the wrongness of indoctrination does not necessarily imply that neurointerventions are impermissible. First, the claim that indoctrination fails to address 'Alex's attitude toward his fellow human beings' implies that indoctrination bypasses one's moral powers in the sense that it seeks to produce certain actions without regard to the motivations that people have for performing those actions. This is not the case with certain neurointerventions, however, which only target precisely a person's attitudes towards others. Second, indoctrination may seek to inculcate a *particular* motivation in someone in order to achieve the right result. For example, Alex's conditioning has ensured that his motivation for acting in the required way is his aversion to feeling ill. Yet this does not seem to be the case with some

neurointerventions, which enable an individual to have access to the *set* of reasonable motivations. The following example makes these differences clearer:

> Suppose there is proof of David's unreasonableness because we find out that he has assaulted transgender people. David is found guilty of the crimes and sentenced to incarceration. Upon his release from prison, there are good reasons to believe that David's views have not changed at all but that he would commit transphobic crimes in the future so long as he thought that he wouldn't get caught. If David is indoctrinated at this point, he will not commit any hate crimes in the future, but he will not identify with his decision not to act unjustly. If asked whether he believes that hate crimes are wrong, he will not be able to justify his (negative) answer in any way that reasonable persons should accept. By contrast, if David is subjected to deep brain stimulation, his empathy could be increased. As a result, he would be able to place himself in other people's shoes and consider things from their point of view, without having his judgment clouded by extreme, irrational emotions, such as hatred. In this case, if asked whether he believes that hate crimes are wrong, David would be able to give appropriate reasons in defence of his (negative) answer.

Now one might respond that both interventions are impermissible because David should fortify his reasonableness on his own, or by being given reasons that he can consider and evaluate. Bublitz and Merkel, for example, seem to be making the second claim when they argue that, by acting directly on the brain, 'direct interventions seem to violate the demands of dignity' and that indirect interventions, on the other hand, such as 'speech or sounds or images, engage with the other's first-person perspective by recognizing and referring to her beliefs and feelings' (Bublitz and Merkel 2014, 73). The thought here is that dignity requires treating persons as agents by respecting their 'first-person perspective', which neurointerventions and any other kind of direct intervention cannot achieve. This claim, however, is quite controversial, given that certain forms of brainwashing that may involve forcing someone to watch or listen to something, for instance, are *indirect* interventions, yet arguably violate the demands of dignity at least as much as direct interventions do. Similarly, there are cases of *direct* interventions, such as certain kinds of psychological

rehabilitation that may involve psychiatric treatment, which seem permissible; and such interventions arguably seem preferable to, say, *indirect* kinds of brainwashing. We may therefore conclude that respecting the ‘first-person perspective’ is neither sufficient nor necessary for the permissibility of an intervention, and that the distinction between direct and indirect interventions only tracks our intuitions regarding the moral and physical boundaries that bodies seem to set, and not the extent to which an intervention shows lack of respect towards human agency. Instead, I have suggested that a more plausible criterion is whether an intervention targets one’s moral powers as a capacity rather than one’s specific beliefs or desires.

To be sure, neurointerventions might not be sufficient as a means for increasing one’s reasonableness to the requisite degree. For instance, many sex offenders do not offend because they have a high sex drive but because they have certain desires that involve harming others and certain views about, for example, women’s worth. In these cases, even if chemical castration prevents further crimes, it might not enhance the offenders’ reasonableness, as they are likely to have the same desires and views post-castration. Of course, this is not the case with all offenders. If offending is due to impulses as contrasted with more stable desires, then giving an individual the capacity to control these impulses can actually increase that person’s autonomy and enable them to act in accordance with their true desires. A good way to distinguish between the two categories of offenders might be by appealing to the difference between an offender’s first-order and second-order desires (Frankfurt 1971, 7). In cases where the first-order desire to violate the requirements of justice is different from someone’s *reasonable* second-order desire, neurointerventions might be sufficient to bring the two into alignment. By contrast, when both one’s first-order and second-order desires are inconsistent with the requirements of justice, it seems likely

that extensive therapy and education will be necessary. In any case, my argument in favour of neurointerventions is compatible with the claim that neurointerventions are not always sufficient means to ensure the required fortification; in fact, it offers reasons in favour of other methods of rehabilitation as well, such as therapy and education.

4.7 Conclusion

In this chapter, I have argued that, as it stands, Rawlsian contractualism seems to exclude unreasonable persons from the constituency to whom justification for the exercise of political power is owed and from the constituency to whom the principles of justice apply. I have suggested that unreasonable persons can be included in the theory once we notice three important reasons why states have a duty to enable them to develop their reasonableness. These reasons are that all persons have a basic interest in developing their two moral powers, that states ought to cultivate reasonableness as a way of promoting justice and that they further have to cultivate reasonableness as a way of promoting stability. My argument suggests that certain policies, such as mandatory education, therapy, and neurointerventions are sometimes required for the state to fulfil its duty to cultivate the reasonableness of unreasonable persons.

Chapter 5

Individuals with Cognitive Disabilities

5.1 Introduction

We have seen that the broader problem of exclusion arises in Rawlsian contractualism due to the theory's commitments to public reason and the significance of social cooperation. The thought is that, because the exercise of political power must be justifiable to *reasonable and rational* persons (who are the only ones that can contribute to society conceived as social cooperation with the purpose of mutual advantage), states and citizens ought to restrict the reasons that they give for laws and policies to the realm of *political* reasons. This in turn means that they ought to refrain from appealing to comprehensive doctrines which include metaphysical accounts of personhood. For example, according to the theory, states and public officials cannot appeal to the idea of psychological continuity in determining who counts as a person in order to justify abortion laws.

Given the constraints set by public reason, a political account of personhood is required. This is because only such an account can be justified in terms that all reasonable and rational persons can be expected to accept in the face of reasonable disagreement on comprehensive doctrines. This account of personhood is provided in the following way: since society is viewed as social cooperation with the purpose of mutual advantage, in order to count as a person, one must have the capacities that are necessary and sufficient to participate in social cooperation. According to Rawlsian contractualism, these capacities are reasonableness and rationality.

This reasoning is what gives rise to the problem of exclusion. So far, I have focused on the exclusion of children, foetuses, and unreasonable criminal offenders—all groups of individuals who seem to have the potential to develop the two moral powers, but only if those who currently count as persons act (or refrain from acting) in certain ways. I have suggested, however, that the question of exclusion similarly applies to two groups – individuals with cognitive disabilities and non-human animals – members of which are usually viewed as lacking the potential to develop the two moral powers to the degree that is required in order to participate in social cooperation. In this chapter, I focus on the first group, namely individuals with severe cognitive disabilities.

On this issue, Rawls notes that

> we do not mean to say, of course, that no one ever suffers from illness and accident; such misfortunes are to be expected in the ordinary course of life, and provision for these contingencies must be made. But given our aim, I put aside for the time being these temporary disabilities and also permanent disabilities or mental disorders so severe as to prevent people from being cooperating members of society in the usual sense.... We may think of these other questions as problems of extension (Rawls 2005, 20).

I suggest, however, that these questions are more serious and threatening to Rawlsian contractualism than Rawls's writings suggest. For, as we have seen, although the theory can easily make provision for contingencies that occur in the lives of individuals who have once been reasonable and rational, this is not the case with individuals who have *never* counted as persons in the political sense, such as individuals who have been born with severe cognitive disability. I argue that the very structure of Rawls's theory of legitimacy and justice gives rise to two related kinds of exclusion.

First, individuals with severe cognitive disabilities are excluded from the constituency of legitimacy. Given that these individuals do not count as sufficiently reasonable and rational, and given that legitimacy requires the justifiability of political power to sufficiently reasonable and rational citizens, they are excluded from the constituency of persons to whom justifiability is owed. This implies that it would be possible that certain policies and laws that might harm individuals with cognitive disability would pass the justifiability test. It might already be objected at this point that being reasonable entails that such laws and policies would never seem justified. However, given that reasonableness tracks the capacity to understand the requirements of justice, this would seem to presuppose too much; for, as we have seen, it cannot merely be assumed that justice issues in requirements with regards to groups that do not count as persons in the political sense.

This brings us to the second question of exclusion, which follows directly from the first: given that individuals with cognitive disability are excluded from the constituency to whom justification for the exercise of political power is owed, they are also excluded from the constituency of justice.[54] More specifically, this means that individuals with cognitive disability are not represented in the constituency that determines the content of justice according to Justice as Fairness; indeed, only sufficiently reasonable and rational citizens are represented in the Original Position. Given that the parties who represent them are mutually disinterested and know that they are reasonable and rational, it seems most likely that they do not have reasons to take the interests of individuals with cognitive disabilities

[54] As I have mentioned in previous chapters, this kind of exclusion is *related* to the one discussed in the previous paragraph because the conditions for inclusion in the justificatory constituency determine the conditions for inclusion in the constituency that agrees to principles of justice. Despite this link, because the former constituency tracks the conditions for legitimacy, while the latter tracks the conditions for justice, it seems to me important for reasons of intellectual clarity to draw a distinction between the two.

into consideration. As a result, the principles of justice that they choose in their pursuit of a fair distribution of the benefits and burdens of social cooperation are likely to favour those that they represent, possibly at the expense of individuals with severe cognitive disabilities. For instance, it has been pointed out that the list of social primary goods that are selected as all-purpose goods that respond to citizens' needs exclude many purposes and needs that are specific to individuals with cognitive disabilities (Hartley 2009, 22). Although these goods, such as income, wealth, and freedom of movement, are important to everyone, there are other goods that are all-purpose goods for individuals with cognitive disabilities and yet are not included in the account of primary goods that is selected by reasonable and rational persons. For instance, resources such as income and negative freedom of movement would not be valuable to individuals with severe cognitive disabilities if they did not translate into actual caregiving. Similarly, Hartley suggests that in order to enjoy the freedom to communicate with others (which non-disabled individuals have free access to), some individuals with disabilities would need access to devices such as voice recognition software (Hartley 2009, 23). And while it is true that primary goods can be interpreted in ways that can accommodate such differences, in practice, the standard interpretation will focus on the average needs of those who are represented in the decision-making constituency, precisely because those included know that they represent sufficiently reasonable and rational individuals (Hartley 2009, 23; Rawls 2005, 178—181).

At this point, it might be argued that, even if individuals with cognitive disabilities could not be included in the constituency that agrees to the principles of justice and, by extension, even if their needs could not be specifically addressed at the level of the designing of the principles of justice, their needs could be met by the principles of justice as they stand. For instance, even if the standard account of primary goods does not include goods that mainly

individuals with cognitive disability need, the principles of justice, such as the liberty and the difference principle, could be interpreted in a way that appeals to these goods. However, as we have seen in previous chapters, it seems questionable whether the principles of justice, as they stand, *can* apply to those who are excluded. For instance, the reason why the principles of justice do not protect non-human animals or foetuses is that the principles of justice are meant to ensure the fair distribution of the benefits and burdens of social cooperation among those who participate in social cooperation; yet these groups cannot participate in social cooperation, as they do not fit the political conception of the person. Thus, as Rawlsian contractualism stands, individuals with cognitive disabilities seem to be excluded from the theory in two (related) ways: first, they are excluded from the constituency of legitimacy and, second, they are excluded from the constituency of justice.

The purpose of this chapter is to respond to these pressing worries, showing how Rawlsian contractualism can recognise in an appropriate way the rights and interests of individuals with cognitive disabilities. I begin by examining the question of whether the problem of exclusion actually applies to all individuals with disabilities, physical and cognitive alike. I then examine a number of solutions that have been offered to this question and show why they fail to produce a more inclusive account of Rawlsian contractualism or a better alternative to Rawls's theory. In the final section of the chapter, I defend a twofold argument that shows how individuals with cognitive disabilities can be included in Rawlsian contractualism without abandoning the theory's commitment to public reason and the role of social cooperation.

5.2 Which Disabilities Raise the Question of Exclusion?

One question that must first be addressed is whether the question of exclusion applies only to individuals who do not pass the threshold of reasonableness and rationality due to severe cognitive disabilities, or, more broadly, to all individuals with physical and cognitive disabilities. If it applies to the latter, then the exclusionary implications of Rawlsian contractualism may be even more serious than one might at first suppose.

Indeed, it has been suggested that the theory's exclusionary implications are even more problematic than it might at first seem due to the theory's emphasis on social cooperation (Badano 2014, Hartley 2009). Given that society is viewed as social cooperation with the purpose of mutual advantage, and given the related emphasis on reciprocity, the argument goes, those who cannot provide sufficient benefits cannot participate in *mutually* advantageous relationships; rather, from the perspective of their cooperators, the benefits that they receive are always greater than their contributions (Badano 2014, 405). So, according to this suggestion, if political personhood ought to be defined by appeal to social cooperation, the criterion of reasonableness and rationality is not sufficient, even if it is necessary. For there are often individuals who *are* reasonable and rational and yet, due to serious physical disabilities, cannot provide society with net benefits (Hartley 2009, 26). This reasoning gives rise to the counterintuitive and deeply exclusionary implication that in order to count as a person one must be not only sufficiently reasonable and rational, but also averagely healthy and with average physical needs.

However, if we accepted this reading of social cooperation, we would have to concede that the same argument applies to differences in talents – that is, native productive capacities – as well. For example, those of us who happen to have fewer talents than others or talents that are viewed as less valuable might potentially benefit from social cooperation more than

we contribute to it. Yet Rawls views differences in talents 'as morally arbitrary and so he maintains that citizens do not have a desert-based claim on what they are able to create, through the use of their talents, in a scheme of cooperation' (Stark 2007, 141; Rawls 1971, 64). Thus, Rawlsian contractualism may face either a problem of consistency (if it treats talents and disabilities differently) or a problem of seriously implausible implications. For, unless this argument – i.e. that it is the capacity to provide positive net benefits that matters – can be refuted, the theory would appear to exclude as a non-person anyone who does not have the capacities, abilities, or talents to ensure a positive net contribution to social cooperation.

I suggest, however, that this argument *can* be refuted, and in two different ways. First, as the example of talents reveals, the net effect of one's contribution to social cooperation depends on the ways in which particular talents fit certain social structures more than others, on the value that the majority happens to attach to different talents, which can differ significantly across societies and time, and on the talents that others happen to have in each society. For example, a society that happened to value classical music would view Beethoven's net contribution as significantly greater than a society that did not attach great value to classical music. In the former society, more individuals would be motivated to participate in economic exchanges with individuals such as Beethoven, which would lead to the production of more primary goods. Similarly, if the former society happened to be populated mainly by individuals with Beethoven's talent, then Beethoven's net contribution would be assessed as less significant than it would be if that talent were as exceptionally rare as it was in fact. According to the social model of disability, disabilities are similarly socially constructed (Shakespeare 2017). Whether or not an impairment affects one's ability to perform certain tasks usually depends on the ways that the tasks are structured. For

example, it is not simply being very short that might prevent someone from working at a restaurant; rather, it is being very short combined with the fact that the environment in which the work takes place has been designed for people of average height.

We can see then that if differences in talents do not give rise to the problem of exclusion, neither do at least some differences in physical and cognitive abilities. The reason is that both instances show that the way society is structured has a great effect on the extent to which someone can participate in social cooperation. Given that theories of justice address the issue of how society *ought* to be structured, we cannot simply assume a given social structure and conclude that some individuals cannot participate in that structure. In fact, this is one reason why the conception of justice as fairness ought to include an appeal to impartiality. The central role of impartiality highlights that the interpretation of Rawlsian contractualism as stressing de facto net contributions to social cooperation is mistaken. Instead, what matters is whether individuals have the capacity to participate in social cooperation in a mutually advantageous way *if* society is structured in a way that allows them to do so. If this is the case, then independently of which social structure will be chosen, everyone who has this potential ought to be included as an equal person. Given that physical disabilities do not generally prevent someone from having this potential if she is reasonable and rational,[55] it follows that the probem of exclusion is not raised by individuals with physical disabilities simply in virtue of their having those disabilities, just as it does not apply to individuals with talents that happen not to be greatly valued. Thus, we can dispute the argument for the broad interpretation of exclusion by stressing that many individuals with disabilities have the capacity to contribute to social cooperation if social cooperation

[55] For example, Stephen Hawking has contributed to social cooperation in ways that most able-bodied citizens have not. To be sure, this would not have been possible if he had not been provided with the technologies that enable him to communicate; yet this would not have been sufficient to conclude that he does not have the potential to contribute to society in significant ways.

is structured in certain ways;[56] what matters is therefore whether they have this capacity and not whether this capacity is actually realised in all social structures.

This is even clearer once we notice that the sense in which individuals must be able to participate in social cooperation does not have to be construed in economic terms. For there are many ways in which reasonable and rational individuals can contribute to society, even if they cannot participate in the labour market and even if their economic benefits outweigh their economic contributions. For example, in the words of Christie Hartley,

> by voting and taking part in policy discussions, they [i.e. individuals] can contribute to the political sphere. These individuals can also make cooperative contributions to the family, an institution that is part of the basic structure. These contributions include providing support and companionship to others, participating in family decision making, and helping children obtain the skills and values they will need as adult citizens. By treating other members of society with respect in civil society, they contribute to building the social bases of self-respect (Hartley 2009, 27).

Economic benefits are one kind of advantage that we receive by participating in social cooperation, but they are certainly not the only one. In order to characterise a person's contribution as advantageous for social cooperation, within the Rawlsian framework, we ask whether she can contribute to the production of primary goods that cooperation produces. These include income and wealth, but they also include other benefits, such as the protection of the basic rights and liberties and the social bases of self-respect. Similarly, economic institutions are only one part of the basic structure, as the latter contains a number of social, economic, and political institutions that are associated with the primary goods that

[56] Of course, there are reasonable feasibility constraints here; for instance, if my potential to contribute to society depends on the availability of a special device that cannot be constructed, then I count as a non-contributor to whom the problem of exclusion applies.

social cooperation aims to produce. Thus, since, as we have seen, according to Rawlsian contractualism, all reasonable and rational individuals have the capacity to make some kinds of cooperative contributions to society that sustain the production of primary goods, for example by upholding various institutions of the well-ordered society via rule-following behaviour, the question of whether they have physical or cognitive disabilities that do not affect this capacity is irrelevant. We can therefore conclude that these disabilities do not raise the question of exclusion.

5.3 Potential Solutions

Even if the question of exclusion does not apply to those who have physical and cognitive disabilities that do not affect their reasonableness and rationality, Rawlsian contractualism must address the problem of exclusion regarding its treatment of individuals with severe cognitive disabilities, understood as these individuals who have not developed the two moral powers to the degree required for social cooperation with the purpose of mutual advantage. Before defending a response to this problem, I will first evaluate a number of competing solutions that have been offered.

5.3.1 *Frontiers of Justice*

Nussbaum's *Frontiers of Justice* is, to my knowledge, the only comprehensive account of the problems that the political conception of the person poses to Rawlsian contractualism. In *Frontiers*, Nussbaum critiques all social contract theories on the basis of their exclusionary effects for individuals with severe cognitive disabilities, non-human animals, and citizens of other nations. Focusing on Rawls's theory, which she views as the strongest contractualist theory, Nussbaum claims that the theory can be defended only if it abandons its commitment to 'the idea of rough equality in power and the related idea of mutual

advantage' (Nussbaum 2010, 77). Regarding the problem of exclusion as it applies to individuals with severe cognitive disabilities, Nussbaum advances an argument from an overlapping consensus which, as we will see, can be interpreted in two ways.

According to this argument, any reasonable conception of justice views individuals with severe cognitive disabilities as full and equal persons and acknowledges that humans are social animals who participate in social cooperation due to their desire to interact with others (Badano 2014, 414). Thus, it is argued that the problem of exclusion that Rawlsian contractualism faces can be resolved by stressing the role that the idea of overlapping consensus ought to play in the theory. This idea is sufficient, according to Nussbaum, to warrant the full inclusion of individuals with severe cognitive disabilities in the constituencies of legitimacy and justice.

This argument can be interpreted in two ways. The first interpretation of the argument is an empirical one and observes that there *exists* an overlapping consensus among reasonable comprehensive doctrines on the claim that individuals with cognitive disabilities ought to count as full, equal persons. If the argument is construed in this way, however, it fails to resolve the problem of exclusion. For it makes the inclusion of individuals with cognitive disabilities *contingent* on whether those who independently count as full free and equal persons happen to think that individuals with cognitive disabilities should count as persons too. Given that individuals with severe cognitive disabilities do not have the capacities that are associated with the political conception of the person, their inclusion depends on whether those who are already included happen to believe that cognitive disability is not sufficient to justify someone's exclusion (Stark 2007, 120). This implies that the inclusion of individuals with severe cognitive disability is contingent on reasonable persons

happening to believe that they should be included; if there happens to be no overlapping consensus on the claim that cognitive disability does not affect someone's moral status, individuals with cognitive disability will still be excluded from the constituencies of legitimacy and justice. That is, they would neither be included in the constituency to whom the exercise of political power must be justifiable, nor would they be included in the constituency to whom the principles of justice apply. We therefore see that using an empirical argument about the existence of an overlapping consensus to justify the inclusion of individuals with cognitive disabilities is objectionably contingent and, it follows, unlikely to resolve the problem of exclusion.

Nussbaum might then respond that the correct interpretation of her argument is a normative one. The claim according to this second interpretation would not be that there *happens* to be an overlapping consensus but that agreeing on the inclusion of individuals with cognitive disability is a *necessary* condition for counting as a reasonable person. That is, her argument would be that a comprehensive doctrine cannot count as reasonable if it does not share a commitment to the inclusion of individuals with cognitive disabilities. Yet, in the absence of further justification, this interpretation of the argument from an overlapping consensus seems question-begging. There are theories of personhood, such as the psychological continuity theory, which imply that individuals who do not experience this continuity are not full and equal persons. Similarly, there are accounts of human nature that dispute the fact that humans are inherently social animals whose motivation for participation in social cooperation is the desire for interaction rather than, say, the desire for their own rational advantage. We cannot merely *assume* that these theories are unreasonable (see McMahan 2002; McMahan 2010; Singer 2010). In fact, disagreement about such theories is precisely the kind of reasonable disagreement on comprehensive

claims that political liberalism is designed to address. Thus, when we are faced with a debate on what ought to count as a reasonable view of personhood, we cannot defend a specific position by stating that this is what all reasonable people think; rather, we must come up with justifications that seem acceptable to others who might hold different reasonable comprehensive doctrines.[57]

Nussbaum might respond that this justification is given by her account of human personality and fulfilment. This account aims to explain why individuals with severe cognitive disabilities ought to be viewed as fully free and equal persons. According to Nussbaum, any reasonable theory of justice must endorse a conception of human personality that places as much emphasis on humans' animal needs as it does on their rational powers (Nussbaum 1995, 83—85; Badano 2014, 414). Satisfying these needs requires, among other things, 'opportunities for sexual satisfaction and choice in reproduction; being able to use the senses, imagine, think, and reason; being able to have attachments to things and persons outside ourselves; being able to form a conception of the good and to engage in critical reflection about the planning of one's own life; being able to have pleasurable experiences; being able to show concern for other human beings; and being able to laugh, to play, to enjoy recreational activities' (Nussbaum 1995, 83—85). On the basis of this account of personhood, Nussbaum makes a twofold claim: all individuals who have at least some of these capabilities count as full and equal persons, and this must be acknowledged by all reasonable comprehensive doctrines. It follows, according to Nussbaum, that all individuals with cognitive disabilities ought to be included in any theory of justice as free and equal persons. Moreover, Nussbaum argues, being reasonable implies

[57] This implies that appealing to intuitions in the context of reasonable disagreement is unlikely to succeed in defending a position that can be the object of overlapping consensus. Indeed, for instance, Peter Singer, whose intuitions are different from Nussbaum's, rejects her appeal to 'intrinsic worth and dignity as just a piece of rhetoric unless it is given some support' (Singer 2010, 337).

acknowledging that humans are social animals who have a commitment to the fulfilment of all persons' needs (Badano 2014, 414). Thus, instead of focusing on whether one has the capacity to contribute to social cooperation, reasonable persons would focus on one's capabilities, thereby ensuring the full inclusion of individuals with severe cognitive disabilities (Nussbaum 2006, 155—223).

The problem with this argument is its perfectionist premises and implications, including a specific account of human nature which views humans as social animals, and the related Aristotelian idea that there are specific capabilities that are necessary for human flourishing (Silvers and Francis 2009, 483). For instance, Nussbaum's argument makes certain assumptions about 'the place of sexual satisfaction, reproductive freedom, strong emotional attachments and even recreational activities in human life', which would certainly contradict a number of reasonable comprehensive doctrines (Badano 2014, 415). Similarly, one can be reasonable – in the sense that one can understand and comply with the requirements of justice – even if one thinks that there is no such thing as human nature. The implicit appeal to perfectionist accounts of human flourishing is clearly inconsistent with Nussbaum's commitment to political liberalism.[58] Moreover, if we have good reasons to endorse political liberalism, as I have assumed that we do, the argument is not only inconsistent with Nussbaum's other commitments; it is also independently implausible.

More importantly for the purposes of this chapter, the argument fails to include individuals with severe cognitive disabilities in the way that Nussbaum intends. Recall that the main reason why the question of exclusion arises in Rawls's theory is the emphasis that is placed

[58] Interestingly, Nussbaum elsewhere endorses arguments against perfectionist liberalism, and offers, as one of her reasons for endorsing capabilities as opposed to functionings, the need to avoid making perfectionist claims (Nussbaum 2011).

on two capacities, namely, the capacity to form, revise, and pursue a conception of the good, and the capacity to understand and comply with the requirements of justice. The problem with regards to disability is that many individuals with severe cognitive disabilities do not possess these capacities. Nussbaum's account, however, appeals to an even greater number of capacities. For example, as mentioned above, her account of human personality includes the capacity to 'use the senses, imagine, think, and reason; to have attachments to things and persons outside ourselves; to form a conception of the good and engage in critical reflection about the planning of one's own life; and to show concern for other human beings' (Nussbaum 1995, 83—85). Presumably, if the two capacities that Rawls mentions have exclusionary effects in terms of some individuals' participation and representation in the constituencies of justice and legitimacy, the fact that Nussbaum similarly appeals to certain capacities makes her argument vulnerable to similar problems.[59]

Nussbaum concedes that some individuals with severe cognitive disabilities would not fit this account of human personality and argues that, in these cases, the political status of these individuals is consistent with paternalistic treatment that would not normally be permissible with fully free and equal persons. In these cases, when people are not capable of using capability spaces, such as a number of rights and freedoms, to make their own choices, Nussbaum concedes that their functionings should be targeted until each person is brought up to 'the minimal level of functioning envisioned for citizens in a just society' (Nussbaum 2006, 190).[60] For example, instead of giving individuals the option of

[59] As we have seen, Nussbaum claims that individuals should have *at least some* of these capacities in order to count as persons, although she does not specify how many or which of these capacities are required and why these capacities would be sufficient.
[60] Briefly, the distinction between 'capabilities' and 'functionings' is that the former track opportunities to achieve well-being, whereas the latter track one's actual achievements in terms of well-being.

exercising their freedom of conscience, states should try to ensure that individuals with severe cognitive disabilities actually enjoy and exercise freedom of conscience. However, this brings us back to the initial questions regarding the question of exclusion. For it is not clear why those who do not fit this account of personhood are to be treated by political liberalism as having an interest in developing the capacities that characterise the political conception of the person in the first place. Nor is it clear why others have a duty to contribute to the excluded parties' development of the relevant characteristics. Moreover, even if sufficient justification were offered for the claim that those who do not count as full persons have interests that give rise to duties, given that Nussbaum's account of interests is not properly *political*, the charge of perfectionism would loom in an even stronger form. This is because Nussbaum's argument would commit states to promoting *functionings*, such as one's attachments to things and persons, which might be inconsistent with a number of reasonable comprehensive doctrines. Thus, Nussbaum's human fulfilment argument fails to provide a political liberal solution to the problem of cognitively disabled individuals' exclusion, as it is perfectionist in a way that makes it inconsistent with political liberalism and still too exclusionary anyway.

5.3.2 Personhood as Minimal Reasonableness and Rationality

A different solution has been proposed by Gabriele Badano, who holds that the problem of exclusion is resolved if we prioritise persons over social cooperation:

> the idea of persons should be regarded as forming the ultimate basis of the common culture of our societies, while social cooperation is aimed at serving the interests of persons. This ordering, which is in line with the liberal common sense that society should serve individuals (and not vice versa), means that no individual can be excluded from membership because of their inability to contribute adequately to social cooperation (Badano 2014, 414).

According to Badano, once we appreciate that persons should be prioritised and that the idea of persons should ground the *common culture* of a well-ordered society, we see that we need a conception of personhood on which there is overlapping consensus – that is, a conception that is consistent with all reasonable comprehensive doctrines. Badano thinks that conception is provided by the Rawlsian account of reasonableness and rationality, stripped of the threshold that is set by the idea of social cooperation with the purpose of mutual advantage. That is, in order to count as a person, one should be '*minimally* reasonable and rational'. Being minimally reasonable is defined as having the capacity to recognise some cases of injustice, such as wrongful violence, as such, while minimal rationality refers to the capacity to have ends or interests, as well as the desire to fulfil them (Badano 2014, 416). In this way, the idea of persons is prioritised over the idea of social cooperation; instead of assessing one's capacity to *contribute* to social cooperation with the purpose of mutual advantage, we ought to focus on whether one fits the idea of persons that is shared by our common culture. Once we adopt this focus, the argument goes, most individuals with severe cognitive disabilities would count as persons and would, therefore, be included in the constituencies of justice and legitimacy.

In order to assess this solution to this problem of exclusion, let us first consider the claim that Rawlsian contractualism prioritises social cooperation over persons. As I argued in the first section of this chapter, the correct interpretation of the role of social cooperation is that, while it is conceived as having the purpose of mutual advantage, what matters for the purposes of determining who counts as a co-operator is whether one has the *potential* to contribute to social cooperation, and not whether one's actual net contribution is positive. This is not, however, because social cooperation is prioritised over persons. Rather, the very reason why social cooperation matters is that it benefits persons by enabling them to

satisfy their fundamental interest in developing and exercising the two moral powers and, relatedly, by giving them access to primary goods, whatever their comprehensive doctrines may be. It is true that the idea of social cooperation in turn shapes the political liberal conception of personhood, as well as the account of politically justified rights and duties. Yet, once again, this does not prove that social cooperation is prioritised over persons. Instead, it shows that, because persons have reasons to participate in social cooperation, and because they face the problem of reasonable disagreement, they acknowledge that the exercise of political power must be justifiable in political, rather than comprehensive terms; and, by definition, political terms implicitly or explicitly appeal to the political idea of social cooperation with the purpose of mutual advantage. We see, then, that even though the idea of social cooperation is central in Rawlsian contractualism, the justification for this role is that social cooperation benefits persons. Thus, the idea of the person takes precedence over the idea of social cooperation, even if the latter shapes the specifically political conception of the former.

Even though it is not true that Rawlsian contractualism prioritises social cooperation over persons, one might still argue that individuals should not 'be excluded from membership because of their inability to contribute adequately to social cooperation' (Badano 2014, 414). In order to explain why they ought to be included, one might claim that the conception of personhood that forms the basis of the common culture of a well-ordered society is one that appeals to minimal reasonableness and rationality, and further claim that all human beings fit this conception of personhood. However, it is not clear how this conception of personhood can be justified in a way that is consistent with political liberalism. As we saw in the previous section, unless a conception of personhood is justified in political terms, typically by appealing to the idea of social cooperation, it is either unjustified or else

justified in comprehensive terms. If it is unjustified, then it is not clear why it ought to be accepted. If it is justified in comprehensive terms, for example by appealing to a metaphysical conception of personhood, it is inconsistent with political liberal commitments. Of course, it might be thought that this is a decisive problem for Rawlsian contractualism. However, since I have assumed that political liberalism is independently plausible, the problem is to be regarded as decisive only if we cannot resolve the problems of exclusion while adhering to the basic tenets of political liberalism, including its political conception of the person. And, as I have stated, I will argue in this chapter that this problem of exclusion *can* be resolved *within* the boundaries of political liberalism.

Now one might object that the conditions of minimal reasonableness and rationality *can* be justified in political terms, by appealing to an overlapping consensus among all reasonable comprehensive doctrines on this account of personhood. Indeed, this seems to be the justification that Badano has in mind when he argues that the three examples of reasonable doctrines that Rawls uses – namely Kant's liberalism, Mill's conception of autonomy, and classic utilitarianism – would all converge on this more inclusive conception of personhood (Badano 2014, 417). As I argued in my discussion of Nussbaum's arguments, however, there are problems with this justificatory approach. First, we might dispute whether there really *can* be an overlapping consensus even among these three doctrines on this idea of personhood. For even if certain aspects of the three doctrines are consistent with this conception of personhood, it seems likely that others are not. And even if there is convergence on an account of *necessary* conditions for personhood, the relevant doctrines might disagree on whether these conditions are *sufficient* for personhood. For instance, the Kantian emphasis of moral agency and autonomy would seem to imply that this conception of personhood is *over*inclusive and that minimal reasonableness and rationality are not,

therefore, sufficient to warrant one's inclusion in the community of full free and equal persons. Furthemore, given that the list of the three doctrines given by Rawls as examples of reasonable doctrines is not exhaustive, there is the further problem that, even if these three doctrines agreed on a conception of personhood, it is likely that at least some others would not overlap with them. Moreover, we have seen that appealing to an overlapping consensus among reasonable and rational persons to justify a conception of personhood is question-begging, for it *assumes* that all positions that do not support this conception count as unreasonable, instead of offering a free-standing justification for that conception. Lastly, as I have already argued, the argument from overlapping consensus aims to include individuals with cognitive disabilities in the constituencies of justice and legitimacy in an objectionably *ad hoc* manner, as it effectively states that one group of persons is affirmed as persons only because another group that is independently defined as persons happens to view them as such in accordance with their comprehensive doctrines. Appealing to an empirical overlap among comprehensive doctrines seems speculative and gives rise to an objectionably contingent condition for the inclusion of individuals with cognitive disabilities.

5.3.3 Reciprocity and Relational Contributions

As I argued in the first chapter, the way in which the idea of social cooperation for mutual advantage is used in Rawlsian contractualism should not be conflated with the way that it is used in contractarian theories such as the one advanced by Hobbes. While the latter focuses on individuals' actual net contributions to society, which are calculated by calculating the benefits that they actually manage to receive,[61] the former appeals to the

[61] Notice that some of these benefits might be gained simply due to inequality in strength or power and that they might be inconsistent with any reasonable conception of justice.

contributions that each person has the potential to make and to the benefits that each person receives consistently with the requirements of justice.

Christie Hartley argues that we can use this distinction between contractualism and contractarianism in order to resolve the problem of exclusion as it applies to individuals with cognitive disabilities. This is because the great majority of individuals with cognitive disabilities do have the capacity to contribute to social cooperation insofar as we understand contributions within the contractualist – rather than the contractarian – framework. So understood, these contributions may take a number of forms. Many individuals with cognitive disabilities are able to make economic contributions either directly, by participating in the labour market provided they are given appropriate training and support, or indirectly, by participating in household labour (Hartley 2009, 28). Moreover, most persons with cognitive disabilities can make relational contributions to society, by participating in relationships (Hartley 2009, 28). The only capacity that is required to make these contributions, according to Hartley, is the capacity for engagement, which is

> the capacity to see another as a responsive, animate being and to recognize the ability of the other to be responsive to something she interprets as a communication to herself. It is the ability to recognize others in a certain kind of way and to attribute to them a certain kind of standing (the standing of a being with whom communication is possible), but it also involves the ability to make some kind of communication to another (Hartley 2009, 28).

This capacity is viewed as crucial because it is both necessary and sufficient for communication, which makes most kinds of cooperation possible. Given that most individuals with cognitive disabilities have the capacity for engagement and, therefore, the capacity to communicate in either verbal or nonverbal ways, they have the capacity to

participate in meaningful relationships and enrich other individuals' lives with love, affection, support, and companionship (Hartley 2009, 29). Since these relationships are a crucial aspect of the social bases of self-respect, it follows that most individuals with cognitive disabilities can contribute to social cooperation in the ways that Rawlsian contractualism requires. Lastly, all individuals with cognitive disabilities make an indirect yet no less important contribution to social cooperation: they help others cultivate virtues, such as humility, kindness, patience, compassion, and the inclination to trust others, that might improve the quality of social relationships and society's stability (Hartley 2009, 29; Silvers and Francis 2005). Thus, according to this line of argument, individuals with cognitive disabilities should be viewed as fully free and equal persons who contribute to society as all fully free and equal persons do. Once this becomes clear, it is argued, the problem of exclusion is resolved; individuals with cognitive disabilities are fully included in the constituencies of justice and legitimacy.

We might doubt, however, whether this approach manages to resolve the problem of exclusion because it seems doubtful that the kinds of contributions that it appeals to would count as contributions in Rawlsian contractualism. More specifically, they would not be viewed as the *right kind* of contributions and in cases where they *are* the appropriate kind, the *amount* of contribution would not be appropriate. Call this issue the question of the *quality* and *quantity* of one's contribution to social cooperation (Badano 2014, 412—3). The quality of a contribution is relevant because some contributions might not offer publicly acceptable reasons for inclusion. Consider, for instance, the claim that someone should be viewed as a full person because she is loved by another person and, according to my religion, provoking feelings of love in others is, in itself, a significant contribution to society. In the face of reasonable pluralism, it is immediately clear that this justification is

not consistent with political liberalism. Similarly, the fact that someone's inclusion might enable individuals to develop a number of virtues might be insufficient to justify that inclusion, unless the cultivation of those virtues is necessary in order to achieve political values. For example, claiming that states ought to cultivate patience, humility, love, and compassion in their already sufficiently reasonable citizens seems perfectionist and therefore inconsistent with the commitments of political liberalism (Badano 2014, 412). Indeed, a number of reasonable citizens hold doctrines that affirm the value of 'emotional self-sufficiency and self-reliance in their relationships with others', which would prevent them from sharing the commitments of a state that seeks to cultivate such virtues (Badano 2014, 413). This shows that individuals with cognitive disabilities may still raise the question of exclusion in Rawlsian contractualism because the quality of their contributions to society implies that they cannot be viewed as full co-operators. The kinds of contributions that they can make to social cooperation are not the kinds that can be viewed as contributions that can be justified in political terms in Rawlsian contractualism.

Similarly, Hartley's suggestion that we focus on reciprocity and individuals' relational contributions does not take into account the *quantity* of these contributions, even if they are the right *kinds* of contributions. For an illustration of the problem of quantity, imagine a person who is minimally reasonable and rational and who can make, say, minimal economic contributions of the right sort but whose economic contributions are *always* considerably less valuable than the benefits she receives. Of course, as we have seen in Section 5.2, it is not the *actual*, but the *possible* quantity of one's contribution that matters. Suppose, however, that there is *no possible* social structure in which that person could make weightier contributions. In that case, it seems clear that, if society is viewed as social cooperation with the purpose of *mutual advantage*, the *quantity* of one's (possible)

contributions matters in order to assess whether that person can contribute in a mutually advantageous way.

Of course, I am not suggesting that those who do not meet the relevant threshold should be excluded. In fact, the purpose of the following section is to show how the problem of exclusion can be resolved. My suggestion at this point is merely that Hartley's appeal to reciprocity and relational contributions does not resolve this problem because it relies on an implausible account of contribution that fails to take the quality or quantity of different contributions into consideration.[62] Once the account of contribution is fine-tuned to these issues, we return to the initial problem of exclusion as it applies to most cases of individuals with severe cognitive disabilities.

5.4 Resolving the Problem of Exclusion

Having examined and rejected the most prominent solutions that have been offered to the problem of exclusion as it applies to individuals with cognitive disabilities in Rawlsian contractualism, I now set out an argument that, it seems to me, succeeds in resolving the problem. This argument is two-pronged. First, I argue, in line with my claims about children and unreasonable criminal offenders, that individuals with cognitive disabilities have an interest in developing the two moral powers; any reasonable theory of justice ought to respect and satisfy this interest. Second, I argue that, even when this interest cannot be satisfied, states and those who currently count as persons have certain duties towards

[62] One issue that arises is how the quantity and quality of different contributions are to be measured. This is an important issue but one that lies beyond the scope of my discussion and one that does not affect my argument, I think. Regardless of the difficulties involved in measuring the quantity and quality of one's contributions, it remains true that, in order to assess one's contributions, one must focus on both the quantity and the quality of these contributions.

individuals with disabilities that spring from a combination of the idea of parental, state, and citizens' duties and the idea of opacity respect.

5.4.1 The Interest in Developing the Two Moral Powers

Recall that in the chapters on children and unreasonable citizens, I argued that all future persons have an interest in developing the two moral powers. This interest is *political*, rather than comprehensive, because it arises from the value of social cooperation. That is, we have an interest in developing the two moral powers because these powers enable us to participate in social cooperation. By participating in social cooperation we receive both extrinsically and intrinsically valuable benefits, which are *politically* justified. The extrinsic benefits are captured by the notion of primary, all-purpose goods, which are valuable whatever our conception of the good may be. The intrinsic benefits are captured by the thought that having the capacity for political autonomy and being an agent of justice is good in itself.

I have argued that, although all those who do not currently count as persons seem to be excluded from the justificatory constituency in Rawlsian contractualism, their interests can in some cases, in fact, be represented. This means that even if they are excluded from *directly participating* in the constituency to whom the exercise of political power is justified and, by extension, in the constituency that designs the principles of justice, their interests can be represented in both constituencies. This can be achieved by requiring those who are included in the constituencies of justice and legitimacy to take future persons' interest in developing the two moral powers into consideration. As we have seen, the main reason why they should be required to do so is to ensure that temporal impartiality is preserved. This is how, for instance, intergenerational justice is incorporated into the

theory: namely, the parties in the Original Position are required to agree to principles on the condition that they would have consented to previous generations following the same principles. Thus, given that those currently reasonable and rational had these interests before they developed the two moral powers, and given our commitment to temporal impartiality, we realise that reasonable and rational persons ought to consider the interest that non-persons have in developing the two moral powers.

This strategy can resolve at least many cases of the problem of exclusion as it applies to individuals with cognitive disability too, because the distribution of the benefits and burdens of social cooperation ought to be consistent with and promote individuals' interest in developing the two moral powers. Given that some individuals with cognitive disability have the potential to develop the two moral powers, it is plausible to say that they also have an interest in doing so. By revising Rawlsian contractualism to ensure that this interest is represented in its theories of justice and legitimacy, the problem of exclusion is, in many cases, resolved (cf. Wong 2010, 129).

Although this argument adequately resolves the problem of exclusion in all cases in which individuals with cognitive disability have the potential to develop the two moral powers, it leaves the problem of exclusion unresolved in cases of severe cognitive disability that prevents one from having this potential. In these cases, it might be argued that one cannot have an interest in developing the two moral powers, because we cannot have an interest in states or things that are not realistically possible. For instance, it would be implausible to claim that I have an interest in being a dolphin, given that it is, and will always be, impossible for me to become a dolphin. Moreover, it would be even more implausible to claim that states and persons have a duty of justice to spend some of their scarce resources

to provide me with the conditions that will enable me to become a dolphin. Thus, the objection goes, if there are individuals with cognitive disabilities who could *never* develop the two moral powers, then acknowledging the interest in developing the two moral powers does not manage to fully resolve the problem of exclusion.

One way to respond to this worry is to point out that it is not possible to predict which individuals have the potential to develop the two moral powers. Indeed, we have ample evidence of this inability, which should make us sceptical of empirical claims that serve to justify exclusion. As Michael Bérubé states in conversation with Peter Singer,

> in the 1920s we were told that people with Down syndrome were incapable of learning to speak; in the 1970s, we were told that people with Down syndrome were incapable of learning how to read. OK, so now the rationale for seeing these people as somewhat less than human is their likely comprehension of Woody Allen films. Twenty years from now we'll be hearing 'sure, they get Woody Allen, but only his early comedies – they completely fail to appreciate the breakthrough of Interiors' (Bérubé 2010, 106).

Thus, given that it is not possible to make reliable predictions about whether or not someone has the potential and therefore the interest in developing the two moral powers, and given all societies' history of grave injustices, states should err on the side of caution and overinclusiveness by treating *all* individuals with cognitive disabilities as possessing the potential to develop the two moral powers (Wong 2010, 141; Hartley 2009, 31). We can conclude that the problem of exclusion as it applies to individuals with cognitive disability is resolved in the way that children's exclusion is resolved: by noticing that many individuals with cognitive disabilities have an interest in developing the two moral powers and that even those who seem to lack this interest should be included because we can never be certain of different individuals' cognitive potential. Therefore, even if individuals with

cognitive disability (and all other excluded groups, such as children) are not included in the constituencies of legitimacy and justice on the terms that reasonable and rational persons are – that is, via full direct participation – they are included in a *different* way; for those who have the capacities that are required for direct participation have a duty to take the main interests of individuals with cognitive disability into consideration in their deliberation on issues of legitimacy and justice.

5.4.2 Parental, State, and Citizens' Duties and Opacity Respect

Even if, however, the problem of exclusion is *partially* resolved by appealing to society's current inability to predict whether one has the potential to develop the two moral powers, it remains pressing under certain conditions, for it leaves open the possibility that exclusion would be justified *if* we knew that someone lacked that potential.

I suggest, however, that apart from the appeal to the interest that some individuals have in developing the two moral powers, there is a second justification for the inclusion of *all* individuals with cognitive disabilities. As with the argument examined in the previous section, this argument does not achieve the full inclusion of individuals with cognitive disability in the constituencies of legitimacy and justice *in the way that reasonable and rational persons are included*. This is because individuals with cognitive disability are not granted full direct participation rights. They are, however, included in a *different* way, as, according to my argument, those who do have participation rights have a duty to take the main interests of individuals with cognitive disability into consideration in their deliberation on issues of legitimacy and justice and come to recognise, as a result, that a number of duties are owed to these individuals.

According to this argument, parents, states, and citizens have a number of duties towards children with severe cognitive disabilities. To explain how the argument works, I begin by exploring the duties that parents, states, and citizens have towards children, setting aside the assumption that these children are cognitively disabled. More specifically, I first explore the duties that parents have towards their children, claiming that they have special responsibility towards them because they are morally responsible for the creation of children's basic interests. I then argue that states and citizens also have certain duties towards children because they benefit from parents' having and rearing children. It may seem, however, that, even if my argument about the duties that parents, states, and citizens have towards children is plausible, children with severe cognitive disabilities constitute an exception; that is, due to the commitments and constraints of Rawlsian contractualism, it might be argued that the justifications that are normally given for the duties that exist towards children are not available for children with severe cognitive disabilities. If this is true, then we are led back to the problem of these individuals' exclusion. I suggest, however, that children with cognitive disabilities *do not* constitute an exception to the broader argument about duties to children, because, when determining what kinds of duties there are towards children, those included in the justificatory constituency ought to show opacity respect towards children. I examine the idea of opacity respect below, but, briefly, this means that those included in the justificatory constituency ought to refrain from assessing whether specific children have the potential to develop the two moral powers.

To examine how the argument works, let us start with the claim that parents have certain duties towards their children. As we have seen, all children have certain interests, such as the interest in developing the two moral powers. This seems to imply that all persons have certain duties towards children qua future persons, just as all persons have certain duties

towards future generations qua future persons. I suggest, however, that parents have special duties towards their own children because they bear special moral responsibility for their existence. The main reason for this is that, according to liberal egalitarian theories of justice, of which Rawls's theory is but one, justice must be sensitive to responsibility (Arneson 1989; Cohen 1989; Dworkin 2000). That is, it must hold individuals responsible for the ends that they have and for the choices that they make on the basis of those ends. In the case of parenting, these theories' sensitivity to responsibility implies that 'parents who choose to have children and who identify with their desire or ambition to have them are justifiably held responsible for the costs of having them' (Olsaretti 2013, 227). Given the costs of having and rearing children,[63] I suggest that the desire to have children can be viewed, *to some extent*, as an expensive taste that some rational individuals form and decide to pursue.[64] As with other expensive tastes, the responsibility that an agent has for forming and pursuing a specific conception of the good gives us a *pro tanto* reason in favour of an unequal distribution of the burdens associated with that conception of the good. This means that, no matter what the interests of children (such as the interest in developing the two moral powers) imply for states and citizens, parents certainly have a special kind of responsibility for fulfilling these interests.

Consider, for example, the case that has been discussed by Ronald Dworkin and G. A. Cohen of someone who happens to only want to eat caviar and champagne. Suppose,

[63] As Olsaretti has argued, there are two broad cases of costs that arise from parents' having children. First, there are 'costs of care', which include all the costs in terms of resources such as money, freedom, or time, that carers have to bear. Second, there are the 'costs of added adult members', which are the costs that will arise once children become adults and join the community of persons – e.g. benefits that might be due to them as a matter of justice (Olsaretti 2013, 229—230).

[64] I restrict this claim by stating that there is a *similarity* between the end of parenthood and other expensive tastes because, as I will argue below, there are important differences too. This suggests that, although parents have special responsibility towards their children, the costs of parenting are not to be borne *solely* by parents.

further, that this expensive taste is actually fundamental to that person's conception of the good, for suppose that this person happens to be unable to get *any* amount of welfare from cheaper goods. Despite the importance that this expensive taste may have in that person's life, we would typically hold that justice requires sensitivity to responsibility, in that people should be held responsible for their ends and the choices they make as a result. Indeed, this is clearly affirmed by Rawls himself, who claims that being a free citizen means that one ought to take 'responsibility for their ends' (Rawls 2005, 33). To see how this argument works in Rawlsian contractualism, notice that in a contractualist framework such as the Original Position it would seem absurd to suppose that reasonable and rational persons would agree to use scarce resources in order to fund that person's expensive tastes. Since there are similarities between parenting and expensive tastes, and since creating new persons also creates new duties towards these persons that have to be distributed (such as the duty to fulfil these persons' interest in developing the two moral powers), we can conclude that the first reason why parents have *special* responsibility towards their children is that parents should be held responsible for the cost of their conception of the good.

At this point, it might be objected that expensive tastes should be assessed as individuals' endowments, not as choices for which they are responsible. The thought here is that expensive tastes lead to inequalities in welfare which are, however, unchosen and uncontrollable (Cohen 1989, 27). It may thus be argued that those who happen to have expensive tastes, such as the desire to have and rear children, are simply unlucky in that their conception of the good happens to be expensive. However, the unchosen character of a taste may suffice to classify it 'as a product of brute luck, but it does not follow that this brute luck is bad as opposed to good or neutral' (Clayton 2000, 76). For example, I may feel that I am lucky for being sufficiently sophisticated to only be content with eating

caviar, despite knowing the cost of caviar. And, although the fact that my taste happens to be more expensive might be a case of bad luck, the crucial point is that consequential responsibility for tastes springs from endorsing them according to our value-judgements, assuming that we have information about the cost of these tastes (Clayton 2000, 74). Consider a pathological liar who regrets having the urge to lie as opposed to someone who consciously and willingly lies. The difference in their second-order volition matters; while the latter might be satisfied with lying due to his upbringing, we would still hold him morally responsible for his actions. Although different aspects of luck are relevant in our assessment of moral responsibility in different cases, luck does not absolve persons from responsibility. Thus, it remains true that parents have *special* responsibility towards their children, which other citizens do not have, precisely because justice ought to hold individuals responsible for their conceptions of the good and the choices they make as a result.

I now turn to the second claim mentioned at the beginning of this section, according to which states and citizens have certain duties towards children. As we have seen, all citizens have certain duties towards children qua future persons. For example, much of my discussion in the chapter on children focuses on defending the duty that all citizens have to enable children to sufficiently develop the two moral powers. This duty, we have seen, exists regardless of the fact that children lack the two moral powers here and now. For, in order to avoid the charge of temporal partiality, those included in the constituencies of justice and legitimacy ought to accept principles on the condition that previous generations would have accepted as well. And given that they would recognise that they had an interest in developing the two moral powers when *they* were children, they would recognise their own duty to enable those who are *currently* children to develop the two moral powers.

There exists, moreover, a second reason why states and citizens have certain duties towards children. This reason is that fulfilling one's parental duties often requires the cooperation of society. At the very least, this gives rise to certain negative duties to refrain from interfering with parents who are discharging their parental duties, which often grounds an indirect justification for certain duties towards children.[65] For instance, in order to ensure that a child will develop self-confidence, it is not sufficient that her parents treat her with respect; it is also necessary that others refrain from treating her disrespectfully. It might seem that this reason for states' and citizens' duties towards children is not justifiable in political liberalism. The thought here would be that the desire to have children belongs to one's comprehensive doctrine and society does not have a duty to enable citizens to fulfil their comprehensive doctrines. If these doctrines give rise to certain duties, the objection goes, these duties must be borne by those who have created them in pursuit of their comprehensive conceptions of the good in the first place.

However, I suggest that the parties in the Original Position would be concerned with whether or not children's interests are satisfied and with creating conditions in which parents can exercise their parental duties because society receives important benefits from parents' having and rearing children. For instance, society has an interest in its own orderly reproduction since a population decline would cause problems such as having sufficient government revenue – say, through taxes that younger citizens would pay – in order to cover government spending for older generations, such as pensions and healthcare. Indeed,

[65] Some thinkers appeal to arguments that are similar to the one I am advancing here in order to argue for the inclusion of individuals with cognitive disabilities. For instance, Eva Kittay implicitly appeals to the idea of parental responsibility in arguing that the parent of a child with severe cognitive disabilities requires their co-citizens' cooperation in order to discharge their parental duties and that this gives rise to non-parental duties towards these children (Kittay 2010, 410).

some have claimed that, for this reason, children can be viewed as public goods, others have claimed that children produce positive externalities, and yet others have claimed that children should be viewed as socialised goods (Casal and Williams 1995, 94; Folbre 1994, 86—90; Goodin 2005; Anderson 1999; Olsaretti 2013). It is beyond the scope of this chapter to question whether children fit the characteristics of each of these category (say, non-rivalry or non-excludability). It is clear, however, that all three kinds of arguments capture an important, plausible claim: that society receives a number of benefits from parents' having and rearing children. Given that society benefits from the creation of children, those included in the justificatory constituency would be concerned with whether or not children's interests are satisfied and would conclude that states and citizens have certain duties towards children. Their reasoning would take two main considerations into account. First, they would want to incentivise citizens to procreate and raise their children in a way that satisfies children's interest in developing the two moral powers. The second consideration would be a consideration of fairness. Namely, given that all citizens benefit from parents' having and rearing children, it would be unfair if parents were required to bear the full costs of childrearing.[66]

We thus see that the second claim I mentioned at the beginning of this section – namely, that states and citizens have certain duties towards children – is true for two main reasons: because states and citizens have duties towards children *qua future adults* and because society often benefits from parents' having and rearing children (which gives rise to duties towards both parents and children).

[66] For a discussion of the principle of fair play with regards to the distribution of duties towards children, see Casal and Williams 1995; Olsaretti 2013.

It might be objected, however, that this argument does not apply to children with severe cognitive disabilities, precisely because they lack the potential to develop the two moral powers. If that is true, then the argument from parents', states' and citizens' duties towards children is conditional on the existence of this potential, which would lead us back to the initial problem of exclusion. For instance, it might be argued that the parties in the Original Position would view children with severe cognitive disabilities as an exception because they would permanently lack the capacity to contribute to society's production of primary goods and because, relatedly, they would lack the interest in developing the two moral powers. For that reason, critics might claim, states and citizens would not have any duties towards these children and it would be up to their parents to decide whether they would like to promote their children's interests.

I shall argue, however, that this objection misfires because, in thinking about the role of parents' and citizens' duties towards children, those included in the justificatory constituency ought to show 'opacity respect' towards children. To show opacity respect is to show 'evaluative abstinence – that is, a refusal to evaluate persons' varying capacities' (Carter 2011, 550). In the case under discussion, this means that those included in the justificatory constituency ought to refrain from assessing children's potential to develop the two moral powers. I contend that there are two main reasons why those included ought to treat all children as opaque in this way.

The first reason is that those in the justificatory constituency, such as the parties in the Original Position, ought to be fair towards all parents. Being fair towards all parents implies that the parties ought to be impartial between parents who happen to be more or less lucky in terms of the resources (including time and effort) that satisfying their parental duties

requires. For instance, suppose that two parties in the Original Position represent two persons who happen to be parents, but only one of them is the parent of a child with severe cognitive disability. If the parties decided to restrict parental and societal duties to cases of non-disabled children, the parent in question would face significant disadvantages in her capacity to pursue her conception of the good. These disadvantages would not be deserved in any way, given that raising children with cognitive disabilities is costlier, and given that having a child whose upbringing is expensive is entirely due to brute luck. Moreover, the parent in question would be facing these disadvantages as a consequence of pursuing *exactly* the same conception of the good as the parent of the non-disabled child – namely, the benefits of parenthood.

At this point, it might be objected that this argument contradicts the claim that all persons must be responsible for their ends and that the desire to have children is yet another comprehensive doctrine that, to some extent, resembles expensive tastes. However, recall that we have already seen two reasons why the parties in the constituencies of justice and legitimacy accept that states and citizens have duties towards children. These reasons were that they have duties towards children *qua future adults* and that society benefits from the creation of children and future generations. Bearing this in mind, we see that the parties in the justificatory constituency have strong reasons to grant that they have duties towards children and future generations. Once they have granted this, they face the additional question of whether they should exclude children with cognitive disabilities, given that these children lack the two moral powers. At that point, however, they would realise that this practice would disadvantage *some* parents who could not have predicted that the reasons that standardly apply in favour of child-rearing and in favour of accepting states' and citizens' duties towards children would not apply to the case of *their* children. In order

to protect themselves against this instance of bad luck, it seems plausible to assume that the parties in the Original Position would decide to insure against the possibility of facing such unfair disadvantages by endorsing a commitment to opacity respect towards all children in their deliberations on whose interests ought to be protected by the state.

Another reason why the parties would conclude that opacity respect ought to be shown to all children is that not doing so would be socially harmful. For if the parties in the Original Position simply decided that there are no duties towards individuals with cognitive disabilities, it is plausible to think that the society of justice as fairness would promote harmful social attitudes that are inconsistent with the political emotions that are necessary for the cultivation of reasonableness. For instance, claiming that even abhorrent acts such as torturing people for fun are morally permissible under some circumstances would potentially cultivate attitudes such as indifference to pain and cruelty. Indeed, as we will see in the next chapter, there is evidence that the behaviour humans have towards animals predisposes them to adopt similar behaviours towards humans (Nussbaum 2006, 329). For instance, individuals who behave cruelly to animals are likely to (want to) behave cruelly to humans too (Stupperich and Strack, 2016; Sereny, 1972; MacDonald, 1963; Hellman and Blackman, 1966). If traits such as empathy and kindness are necessary emotions for reasonableness – the capacity to understand and comply with the requirements of justice – then these emotions should be cultivated. Given that cultivating reasonableness is a complex process that requires nurturing certain emotions, we cannot assume that citizens could in practice view torturing some humans for fun as morally innocuous and yet otherwise remain reasonable. So we have a second reason why the parties in the Original Position would conclude that a commitment to opacity respect is required: they would want

those that they represent to develop the traits and emotions that are necessary for or at least strongly correlated with reasonableness.

Another example of how assessing children's potential for developing the two moral powers and excluding those who lack the potential from the realm of justice could be socially harmful is that this exclusion would disincentivise many parents from having children, for they would fear that if they happened to have children with severe disabilities, they would have to carry the full burden of child-rearing, such as providing healthcare, without any kind of social support. Given the benefits that society receives from parents' having and rearing children, treating all children as opaque when determining whether parents, states, and citizens have duties towards them would be socially beneficial.

At this point, it might be objected that the idea of opacity respect would imply that different individuals would have to be treated in the same way. If that were the case, it would be impossible to justify different kinds of education or meeting specific medical needs that individuals with severe cognitive disability would have without violating the requirements of opacity respect. For, in order to assess these different needs to begin with, states would have to assess those individuals' moral powers. I suggest, however, that the idea of opacity respect does not imply that there should never be any judgments about different individuals' capacities when determining the policy implications that follow from specific duties to them. For example, when it comes to interpreting the policy implications of the equality of opportunity principle, the physical disabilities or the parenting that someone has happened to have received may be relevant in order to determine what the principle requires for the treatment of *that* person. Similarly, the duty to enable a child to develop her moral powers may have different implications for different children who face different

hurdles in their development of the two moral powers. Rather, the pertinent point is that when it comes to determining whose interests should be addressed at all, the precise development of a child's moral powers is irrelevant. As we have seen, those in the justificatory constituency would agree on this claim in order to ensure both fairness to all parents and that they avoid socially harmful consequences, such as disinsentivising potential parents from procreating.

In this way, the argument from parents', states', and citizens' duties towards children and the requirement of opacity respect succeeds in resolving the problem of exclusion as it applies to individuals with severe cognitive disabilities; in other words, individuals with severe cognitive disabilities are not an exception to the argument about duties to children because those included in the constituencies of justice and legitimacy ought to treat all children as opaque.

If individuals with cognitive disabilities are not an exception to the argument about duties to children, they can be included in Rawlsian contractualism in the same way that non-disabled children are. Although they are not fully included in the way that reasonable and rational persons are (for example, they do not have participation rights in the justificatory constituency, just as children and future generations do not), those who are fully included ought to take their basic interests into consideration. In the case of non-disabled children, these interests are the interests they have *qua future persons* (in the political sense). Although *some* individuals with cognitive disability will never satisfy the conditions of political personhood, we have seen that there are reasons to treat all children as if they will. This implies that those who do count as persons here and now ought to refrain from treating

individuals with cognitive disabilities as an excluded group; that is, they ought to take their main basic interests into consideration as they ought to do with non-disabled children.

5.5 Conclusion

In this chapter, I have examined and rejected a number of solutions to the problem of exclusion as it applies to individuals with severe cognitive disabilities. I have argued that although these suggestions fail to resolve the problem, Rawlsian contractualism can be protected by virtue of a twofold argument. First, I have claimed that the solution to the problem of children's exclusion can, in fact, apply to a large number of cases of cognitive disability. That is, many individuals with cognitive disabilities have an interest in developing the two moral powers, which ought to be taken into consideration by the members of the constituencies of justice and legitimacy. Whenever it is not clear whether one does have the potential and therefore the interest in developing the two moral powers, states and citizens should err on the side of caution and treat these individuals as they would if they did have this interest. Moreover, I have argued that there is a second reason why Rawlsian contractualism does, in fact, include individuals with cognitive disabilities, which does not depend on whether or not they have the potential to develop the two moral powers. This reason is that creating children gives rise to special parental duties and related societal duties towards children. These duties apply to children with severe cognitive disabilities because those included in the justificatory constituency – e.g. the Original Position – ought to show opacity respect to all children and parents, and the only way to do so is to refrain from assessing a child's moral powers when it comes to the question of whose basic interests ought to be protected.

Chapter 6

Non-Human Animals

6.1 Introduction

Having examined how the question of exclusion arises with regards to individuals with severe cognitive disability, I now turn to another group that lacks the potential to develop the two moral powers: non-human animals. In *A Theory of Justice*, Rawls acknowledges that his contractualist theory of justice inevitably excludes animals and that we need to turn to metaphysics to account for our moral duties towards animals:

> a conception of justice is but one part of a moral view. While I have not maintained that the capacity for a sense of justice is necessary in order to be owed the duties of justice, it does seem that we are not required to give strict justice anyway to creatures lacking this capacity. But it does not follow that there are no requirements at all in regard to them, nor in our relations with the natural order. Certainly it is wrong to be cruel to animals and the destruction of a whole species can be a great evil. The capacity for feelings of pleasure and pain and for the forms of life of which animals are capable clearly imposes duties of compassion and humanity in their case. I shall not attempt to explain these considered beliefs. They are outside the scope of the theory of justice, and it does not seem possible to extend the contract doctrine so as to include them in a natural way. A correct conception of our relations to animals and to nature would seem to depend upon a theory of the natural order and our place in it. One of the tasks of metaphysics is to work out a view of the world which is suited for this purpose; it should identify and systematize the truths decisive for these questions. How far justice as fairness will have to be revised to fit into this larger theory it is impossible to say. But it seems reasonable to hope that if it is sound as an account of justice among persons, it cannot be too far wrong when these broader relationships are taken into consideration (Rawls 1971, 448—9).

However, once we take the requirements of political liberalism into consideration, we see that Rawlsian contractualism faces a serious problem of exclusion with regards to non-human animals. This is because, due to the constraints of public reason, it is not permissible for states to appeal to metaphysical or other comprehensive doctrines. As a result, even if some citizens recognise that they have certain duties towards animals, these duties cannot be recognised on an institutional level, nor can they be enforced by states. As we will see, this can lead to extremely counterintuitive implications, such as the view that states should not intervene when citizens grievously mistreat animals.

The question is whether Rawlsian contractualism really does face this serious problem of exclusion and whether the family of liberal conceptions of justice that are consistent with the requirements of public reason cannot include an account of animal rights. As we have seen in the preceding chapters, the broader question of exclusion arises because Rawlsian contractualism presupposes that those who decide on the legitimacy and justice of laws and principles are those for whom those laws and principles are constructed: reasonable and rational individuals who have the capacity – due to their reasonableness and rationality – to participate in society construed as social cooperation with the purpose of mutual advantage (Nussbaum 2006, 349). As I have argued, this raises questions regarding the exclusion of different groups who – temporarily or permanently – lack the two moral powers.

We have seen that the way in which this issue is resolved in the case of children, foetuses, future generations, unreasonable citizens, and some individuals with cognitive disabilities is through the argument from temporal impartiality. The thought is that these individuals have an interest in developing the two moral powers and that the parties in the Original

Position should take this interest into consideration because they should not demonstrate temporal partiality. There are two groups to whom this argument does not apply: individuals with cognitive disabilities who lack the potential to develop the two moral powers and non-human animals (who similarly lack this potential). The inclusion of the first group is warranted by the argument from parental, state, and citizens' duties and opacity respect; briefly, the idea is that all parents, states, and citizens have duties towards children for a number of reasons. Children with cognitive disabilities do not constitute an exception to the broader argument about duties to children, because, when determining which children's interests ought to be protected, those included in the justificatory constituency ought to show opacity respect towards children, by refraining from assessing their moral powers. This argument, however, does not apply to the case of non-human animals. And given that it is not possible to appeal to metaphysics in order to justify their inclusion, the question of exclusion with regards to animals remains pressing.

According to some philosophers, the exclusion of animals is not a problematic feature of Rawlsian contractualism; in fact, it may even be a virtue of the theory. For instance, according to Peter Carruthers, one of the virtues of contractualism is that it can account for the intuitively plausible view that there are significant moral differences between human beings and non-human animals. Given that all humans have a strong attachment to their family members, Carruthers argues, regardless of whether they are infants, elderly, or individuals with severe cognitive disabilities, the interests of these individuals would be indirectly represented in a contractualist situation such as the Original Position (Carruthers 1992, 5—6). The attachment that many humans feel to animals, by contrast, is not 'a normal product of human emotional mechanisms, but out of (what they take to be justified) moral beliefs' (Carruthers 1992, 8—9). And precisely because contractualism aims at the

fair and stable social cooperation of all citizens, the moral rules that it supports must be 'psychologically supportable'; that is, given what we know about human psychology, these rules must be naturally supportable by rational agents (Swanson 2011, 4). Since the main aim of contractualism is the construction of justified moral beliefs, we cannot *presuppose* the validity of moral beliefs, such as the belief that animals possess moral status and rights. It follows, claims Carruthers, that the reasons why those represented in a contractualist situation such as the Original Position would indirectly represent the interests of other humans who may be excluded would not apply to the case of animals. In this way, contractualism can draw a significant distinction between humans and non-human animals and this distinction is, according to Carruthers, so intuitively plausible that it actually makes contractualism even more attractive.

There is a number of problems with this argument, however. First, it is not clear that *all* humans feel attached to their family members regardless of their capacities. Similarly, we can argue that most humans do feel some kind of attachment to animals, especially with regards to their pets (if they happen to have pets) and some humans would liken this attachment to the attachment they feel to family members.[67] Third, we can question whether emotional attachments that some humans happen to have should be given such great importance, at least in determining principles of justice. For example, we often tend to prioritise the interests of our loved ones – yet this kind of partiality is exactly what the

[67] Carruthers does consider this point, in response to which he claims that this attachment is *not* a universal part of human nature. Moreover, even when this attachment is present, he argues, no human would think that their pet is of equal moral worth to any human. In order to prove this point, he claims that we would give a dog away if our child were allergic to it, just as we would give a carpet away (Swanson 2011, 6). This response is not convincing, however. For imagine what most parents would do if their child had an allergy to another human or, more realistically, if another human had an incurable infectious disease that threatened that child. Presumably, most would want to keep their child away from that human, just as they would in the dog or the carpet case. Yet it would be implausible to deduce from this that the person in question does not have equal moral status.

Veil of Ignorance is designed to preclude. Thus, it would seem strange to allow such reasoning when considering the issue of whether the exclusion of animals in contractualism about justice is problematic or not. It is clear, then, that neither Carruthers' argument for the exclusion of animals nor his view that this exclusion would be a merit of contractualism hold their ground.

To see why the exclusion of non-human animals *is* problematic, let us return to the three problems that, I have argued, the broader question of exclusion creates for Rawlsian contractualism. Recall that these problems were temporal partiality, indeterminacy, and undesirable implications. Although the first does not apply to the case of non-human animals, and the second has been addressed in previous chapters (as I argued, the theory is not really indeterminate, even if it is inconclusive, but inconclusiveness is not a worrying issue), the third problem – the problem of undesirable implications – remains pressing. This is because the exclusion of animals would imply that there are no positive or negative enforceable duties towards them; if that is true, then states and individuals qua citizens would not be able to publicly claim that even morally reprehensible acts like torturing animals for fun are impermissible.[68] For, as we have seen, if non-human animals are entirely excluded from the realms of justice and legitimacy, then there cannot be *any* enforceable duties towards them, nor is it clear why their interests would be taken into consideration by those who *are* included. The counterintuitiveness of this implication would create a problem for any theory, but even more so for a theory such as Rawlsian contractualism that relies on the method of reflective equilibrium. For, although intuitions are not sacrosanct according to that method, they ought to be taken seriously when reaching

[68] This implication of Rawlsian contractualism has been examined by a number of theorists, most of whom conclude that, at the very least, this exclusion presents a serious challenge to Rawls's theory. See for example, Bernstein 1997; Plunkett 2016; Tom Regan 1988, 163—74; Elliot 1984; Garner 2003a.

a considered judgment – especially when they are as strongly held as the intuition that torturing an animal for fun is wrong.

Some have suggested that there are further reasons why Rawlsian contractualism is threatened by the exclusion of animals. First, according to Peter Singer, favouring humans over animals is a kind of 'speciesism' that is similar to other 'isms' that we view as morally problematic, such as racism and sexism (Singer 2010, 336; Plunkett 2016, 7). Indeed, unjustified bias *between* different species may be analogous to unjustified bias *within* a species and justifications for the former may be used for the latter. For instance, Bernard Williams's argument that 'since we humans are doing the judging, we are entitled to prefer our own kind' seems to offer a justification that could be used for morally reprehensible views such as racism or sexism (Williams 2008; Singer 2010, 336). Thus, according to this argument, the exclusion of non-human animals from Rawlsian contractualism makes the theory vulnerable to the charge that it shows unjustified bias in favour of humans which is akin to the unjustified bias that is shown in cases of discrimination *within* the human species.

But if the problem is *unjustified* bias, then *if* one could argue that the exclusion of animals is justified, the analogy between favouring humans and other kinds of '-isms' would collapse. It seems that Rawlsian contractualism does exactly that: it justifies the exclusion of animals by noting that justice concerns the fair distribution of the benefits and burdens of social cooperation. Because non-human animals do not and cannot participate in social cooperation, they are not to be included in a theory of justice.[69] Thus, because Rawlsian

[69] Some might dispute this claim and argue that non-human animals can, in fact, participate in social cooperation. I address this argument in Section 6.2.2. I here rely on the assumption that non-human animals lack this capacity, as is widely supposed.

contractualism does offer a justification for the exclusion of animals, we might think that it cannot straightforwardly be accused of objectionable speciesism, even though, as I have argued, that exclusion is problematic for a different reason – namely, due to its deeply counterintuitive implications.

A further reason that has been given as to why the theory is threatened by its exclusion of animals is that this exclusion reveals a cultural bias: according to Bernstein, 'this is because the parties in the Original Position assume that they cannot transform or be reincarnated into animals and therefore they don't consider animal rights (and this is an Eastern, not a Western idea)' (Bernstein 1997, 60). This objection, however, fails to appreciate that these ideas are excluded not because they are Eastern but because they appeal to comprehensive doctrines that are themselves excluded due to the requirements of public reason. In fact, precisely because societies are characterised by reasonable disagreement, the ideal of public reason is a way of respecting cultural diversity. Similarly, Rawlsian contractualism does not appeal to the political conception of the person because it is a Western idea; rather, this conception *arises* from the contractualist idea of viewing society as social cooperation with the purpose of mutual advantage, which is not obviously a distinctively Western idea. All that matters, therefore, is whether one's inclusion can be justified in political terms. In this case, animals are not excluded due to an assumption that humans cannot be reincarnated into animals; rather, they are excluded because (whether or not they have been reincarnated), they currently lack the two moral powers and cannot participate in social cooperation. Similarly, even if one could be reincarnated into a pebble, pebbles are nevertheless to be excluded because their inclusion cannot be justified in political terms.

But perhaps we could put the objection in different terms. The worry here would not be that the exclusion of animals *reveals* a cultural bias; rather, the worry would be that this exclusion gives rise to consequences that might *foster* a cultural bias. That is, perhaps the objection shows that excluding animals from the realm of justice is inconsistent with a number of comprehensive doctrines which view animals as sacred and believe in reincarnation, thus leading to the exclusion or marginalisation of a number of comprehensive doctrines. Thus, the critic would conclude, certain comprehensive doctrines such as Eastern ideas about reincarnation might not flourish. Yet, as we have seen, one can hold that humans can be reincarnated into animals and still hold that animals are to be excluded from the realm of justice and legitimacy because they are currently animals. Moreover, even if it were true that some conceptions of the good would not flourish, it is not clear why this is an objection that threatens Rawlsian contractualism. For instance, it is conceivable that some anti-abortion religious doctrines would not flourish in a well-ordered society that promoted pro-choice attitudes, and yet it is not clear why that would constitute an objection to Rawlsian contractualism. Given its antiperfectionist commitments, the theory is opposed to promoting *specific* conceptions of the good. And if, as I have assumed, this commitment is justified, then it is far from clear why states should worry about whether specific conceptions of the good happen to flourish. More generally, we see that, even though the exclusion of animals is objectionable for a number of reasons, these reasons do not include the worry that this exclusion reveals or encourages a cultural bias.

Given that the question of animals' exclusion gives rise to a number of problems for Rawlsian contractualism, responding to this question remains a pressing issue for the theory. In pursuit of this task, the chapter takes the following form. In the first part of the chapter, I examine and reject solutions that have already been offered in response to the

question of the exclusion of animals. These solutions are: an argument from the intuitive equality of humans and animals; the claim that we have relationships with animals for which the two moral powers are not required; the argument that contractualism could include an appeal to altruism; an argument from overlapping consensus; an argument from species membership and average species capabilities; and an intuitive argument in favour of trusteeship. After rejecting these arguments as potential solutions to the problem of exclusion, I offer an alternative solution in the second part of the chapter. I argue that there are duties of compassion that establish a *pro tanto* reason to take animals' basic interests into consideration in the justificatory constituency's deliberation. These duties exist because, given our understanding of human psychology, the cultivation of citizens' reasonableness requires the cultivation of a concern for all sentient beings' basic interests. Although these duties of compassion are less weighty than the duties that I have defended towards other groups that may seem to be excluded from Rawlsian contractualism, they nonetheless address the question of exclusion in a way that prevents the theory from having extremely counterintuitive implications.

6.2 Potential Solutions

6.2.1 The Intuitive Equality Argument

One solution that has been offered to the problem of exclusion in Rawlsian contractualism can be described as the 'intuitive equality' argument. The thought is that 'if a property is *undeserved* in the sense that its possessor has done nothing to merit its possession, then its possessor is not morally entitled to whatever benefits accrue from that possession' (Rowlands 1997, 238—9). One's membership in a specific species and one's possession of reasonableness and rationality are entirely undeserved properties. Therefore, the correct

understanding of Rawls's contractualism, the argument goes, should lead us to the conclusion that animals should be represented in the justificatory process and that the principles of justice should fully apply to them, as they apply to any reasonable and rational person (Rowlands 1997; Rowlands 2013; Van DeVeer 1979, 372—73, 374; Elliot 1984,104—5; Regan 1988, 171—74, 193).

However, there are two problems with this argument. The first is that the argument is overinclusive. This is because, if we follow this line of reasoning, we will be led to the conclusion that any organism, including plants, ought to be included in the constituencies of justice and legitimacy. For, just as in the case of animals, the fact that plants lack the two moral powers is entirely undeserved. It might therefore seem that it is unfair to exclude plants from the realm of justice and legitimacy. Yet this implication is so counterintuitive that it renders the intuitive equality argument implausible.

In defence of the intuitive equality argument, someone might dispute the overinclusiveness objection. That is, someone might argue that the intuitive equality argument would not serve to include plants because it would not make sense for the parties in the Original Position to worry about being a plant. It seems that the relevant consideration here is whether a being has certain features that make it plausible and possible for humans to imagine what it would be like to be in their place. Since plants are not sentient, the intuitive equality argument does not apply to them (Rowlands 1997, 245).

However, in order to include someone in accordance with the intuitive equality argument, it is not clear why it is necessary for the parties to be able to put themselves in that being's shoes. All that matters for inclusion, according to this argument, is that a being has certain

properties that give rise to interests and that these properties are undeserved. Indeed, in order to represent humans, the parties need not have the capacity to imagine what being a human is like. Instead, they ought to have a good understanding of what interests reasonable and rational persons have. Similarly, however, we notice that the parties can understand the interests that plants have, even if they lack the capacity to imagine what being a plant is like. For instance, given that plants have neurons and are living organisms, we can plausibly think that they have an interest in being watered.

Thus, we may question whether the restriction of imagining what it is like to be someone else is plausible to begin with, given that it seems objectionably anthropocentric. The mere fact that it is difficult for humans to imagine what being a plant is like does not show that plants do not have interests. In fact, the anthropocentricism of this argument is inconsistent with the premises of the intuitive argument from equality; for whether or not one happens to have the capacities that humans have, and therefore the capacities that it is easy for humans to imagine having, is entirely due to luck. Therefore, the proponent of the intuitive equality argument cannot respond to the overinclusiveness objection by appealing to our inability to imagine what being a plant is life. The overinclusiveness objection to the intuitive equality argument stands.

The second problem with the intuitive equality argument is that it ignores that Rawlsian contractualism views justice in political constructivist terms. It is 'constructed *by* human beings, in order to facilitate interactions *between* human beings, and in order to make possible a life of co-operative community' (Carruthers 1992, 102). This follows from the foundational idea of society as social cooperation with the purpose of mutual advantage. Therefore, in the absence of further justification, we cannot accept that the intuitive

equality argument would apply to those who permanently lack the capacity to participate in social cooperation without threatening the internal coherence of the theory. Even if we are sympathetic to the broad thrust of the intuitive equality argument, it does not provide us with a reason to include animals in the realms of justice and legitimacy. For, given that who is represented is determined by the way in which society and its purpose are conceived, the intuitive equality argument can apply only to those who are to be represented. This means that undeserved inequalities are viewed as morally arbitrary and give rise to duties of justice insofar as they concern those who are already included in the constituencies of justice and legitimacy. It follows that the intuitive equality argument is not sufficient to warrant the inclusion of animals unless we presuppose that animals ought to be represented; yet, as animals do not possess the potential to develop the two moral powers, it is far from clear that this assumption may be made. The intuitive equality argument ought to be restricted to human beings because widening its scope would be unjustified and would threaten the coherence of the theory. So the intuitive equality argument is not consistent with the political constructivist commitments of Rawlsian contractualism and cannot, therefore, enable Rawlsians to resolve the problem of exclusion. As we have seen, however, despite the theory's political constructivist tenets, the question of animals' exclusion creates a serious problem for Rawlsian contractualism, for it leaves the theory vulnerable to the problem of unacceptable implications. Rawlsians must then look elsewhere for a solution to the problem of animals' exclusion.

6.2.2 Human-Animal Relationships

A different solution to the problem of exclusion in Rawlsian contractualism is found in the suggestion that we already have relationships with animals, which are effectively a form of cooperation. Recall that the reason why reasonableness and rationality are required for

inclusion is that they are viewed as necessary conditions for one's participation in social cooperation. This argument, however, disputes this claim by noting that we *do* cooperate with animals. If this is true, then the two moral powers are not necessary for inclusion; all that matters is whether a being can form cooperative relationships with others and, according to the proponents of this view, animals effectively can and do do that. In the words of Nussbaum,

> we have many types of relationships with members of other species, relationships involving responsiveness, sympathy, joy in excellence, and concerned interaction, as well as manipulativeness, indifference, and cruelty. It seems plausible to think that these relationships ought to be regulated by justice (Nussbaum 2006, 326).

This is meant to show that full reasonableness and rationality are not, in fact, required for social cooperation; the only condition for inclusion is the capacity to participate in relationships (Tucker and MacDonald 2004; Swanson 2011; Smith 2008, 206—7). Consequently, it is argued, all beings who can form relationships, including non-human animals, ought to be viewed as persons to whom the principles of justice apply.

Recall, however, the problem that we have already seen with regards to this suggestion in the previous chapter: this suggestion ignores that both the quantity and the quality of one's contributions to social cooperation matter in order to characterise these relationships as truly reciprocal and cooperative (Badano 2014, 412—3). The *quality* of a contribution matters in the sense that it must be of the *right kind*; for some contributions might not offer publicly justifiable reasons for one's inclusion. Consider, for instance, the claim that some of the contributions that animals can make are valuable because they are viewed as such by my religion. Given the existence of reasonable pluralism, this justification for inclusion would not pass the test of public reason. Indeed, this is why the metric of one's contribution is

given by the idea of primary goods. It is difficult to see how the contributions that animals are capable of making would pass this test and, as a result, their minimal reasonableness does not seem sufficient to grant them membership in the constituencies of justice and legitimacy. We therefore see that the kinds of contributions that animals can make to social cooperation are not the kinds that are *publicly* justifiable in Rawlsian contractualism.

Similarly, it seems doubtful that the *quantity* of the contributions that non-human animals are capable of making would be sufficient to ground a cooperative relationship. As we have seen in the chapter on cognitive disability, in order to view the problem of quantity we may imagine a person who is minimally reasonable and rational and who can make, say, minimal economic contributions which are of the right sort but which are *always* significantly less valuable than the benefits that person receives. We may further suppose that there is no possible social structure in which that person could make contributions of greater quantity. In this scenario, it seems clear that, since society is viewed as social cooperation *with the purpose of mutual advantage*, the quantity of that person's contributions would not be assessed as mutually advantageous in any possible social structure.

Therefore, even if and when it is true that a non-human animal has the capacity to participate in relationships with humans, that mere fact is not sufficient to justify the conclusion that these relationships count as morally relevant contributions to social cooperation. Thus, it seems impossible to justify a duty to include animals as full and equal persons in Rawlsian contractualism on the basis of their contribution to social cooperation. More broadly, appealing to existing or possible relationships between humans and non-human animals does not succeed in resolving the problem of exclusion.

6.2.3 Altruistic Contractualism

An alternative suggestion that has been offered by Mark Bernstein is to stipulate that the parties in the Original Position are not mutually disinterested but, rather, that they are motivated by altruistic considerations (Bernstein 1997, 66). Driven by such considerations, the argument goes, the parties would concede that animals' interests matter as much as humans' interests and agree to principles of justice that reflect this equality. This would resolve not only the problem of animals' exclusion, but the problem of exclusion as it applies to any being that can be said to have interests.

Now it might be objected that the assumption of altruistic motives is inconsistent with the assumption that we are attempting to formulate principles of justice for the circumstances of justice – that is, circumstances that make 'human cooperation both possible and *necessary*' (Rawls 1971, 126, my emphasis). In the words of Hume,

> 'tis only from the selfishness and confin'd generosity of men, along with the scanty provision nature has made for his wants, that justice derives its origin (Hume 2000, 495).

According to this objection, if everyone were altruistic, there would be no reason to even talk about the fair distribution of resources. Reasonable and rational citizens and those who represent them in the Original Position would not need to bargain and agree on principles of justice because citizens would always be driven by considerations about every being's well-being.

Bernstein considers this objection and argues that it relies on the controversial assumption that there is such a thing as human nature that is inalterable. But once we realise that most, if not all, of our traits are malleable, Bernstein argues, we can imagine that the contractors

would agree on principles of justice that bind them in case some of those they represent eventually act self-interestedly (Bernstein 1997, 68).

Yet this response highlights that the argument in favour of an altruistic contractualism does not rest on the claim that humans *are* altruistic but that they ought to be. For if there is no such thing as human nature, and if human nature is *malleable*, then the argument cannot be that humans *are* altruistic in their nature. Rather, the argument would necessarily be that human personality must be *shaped* so that it becomes altruistic. But why should we *assume* that humans have an enforceable duty to be altruistic unless this is required by a theory of justice? Notice that the answer to this question cannot be that this assumption is required by justice because that would be circular; we would be using an appeal to altruism to show what justice requires and we would then be claiming that there is a duty to be altruistic *because* justice requires it. Thus, the objection that an altruistic contractualism would be inconsistent with the assumption about the circumstances of justice stands.

There is, moreover, an even more pressing problem with appealing to altruism to resolve the question of animals' exclusion: the stipulation that the parties in the justificatory constituency are altruistic rather than mutually disinterested is not consistent with the demands of political liberalism because it relies on a comprehensive doctrine about the significance of altruism. Given that there can be reasonable citizens who comply with the requirements of justice yet who are not *always* or not *necessarily* driven by altruism, appealing to altruism in the justificatory process would violate the requirements of public reason. And this is the case even if there is broad agreement on the claim that altruism is generally good and that it should be promoted and cultivated. For even if there is overlapping consensus on the goodness of altruism, there is reasonable disagreement on

whether enforceable duties can arise from it. Consider, for instance, that most, if not all, reasonable citizens would agree that it is admirable when someone donates a kidney out of altruistic motivations. Despite this agreement, however, it does not follow that there is an enforceable duty to donate one's kidney, nor should we expect consensus on this implication. Unless we presuppose that altruism is part of reasonableness (which would be an extremely controversial claim, given that it is difficult to see why justice requires altruism), endorsing an altruistic contractualism would be akin to endorsing a comprehensively justified contractualism. But, as we have seen, doing so would violate the demands of public reason. It is therefore clear that Rawlsian contractualism cannot be recast as an altruistic contractualism and that, consequently, the problem of animals' exclusion remains.

6.2.4 Appealing to an Overlapping Consensus

As we saw in the previous chapter, Martha Nussbaum attempts to resolve the question of exclusion in the case of individuals with severe cognitive disability partly through the idea of an overlapping consensus. She makes the same argument with respect to the exclusion of animals. According to this argument, all that is needed in order to include animals is the existence of an overlapping consensus on the view that animals are free beings with moral status. In order to respect their citizens' overlapping consensus, states would then be required to publicly recognise that animals have strong interests and rights that give rise to duties of justice (Nussbaum 2006).

The problem is that there is not, in fact, overlapping consensus on whether animals do have full moral status and rights. One could respond that what matters is that there *could* be overlapping consensus on this idea. If that were a possibility, then perhaps states should try

to foster that view. But as we have seen in the chapter on cognitive disability, appealing to the idea of an overlapping consensus in order to address the question of who counts as a rights-bearer is question-begging. There are theories of personhood which imply that animals do *not* have full moral status, and we cannot merely *assume* that these theories are unreasonable.[70]

6.2.5 Capabilities and Species Membership

Nussbaum's second argument for the inclusion of animals appeals to the ideas of capabilities and species membership. According to this argument, each being should flourish as the sort of being that it is. Any being 'born into a species has the dignity relevant to that species, whether or not it seems to have the "basic capabilities" relevant to that species. For that reason, it should have all the capabilities[71] relevant to the species, either individually or through guardianship' (Nussbaum 2006, 347). States should then ensure the dignity and flourishing of all beings – including animals – in accordance with the capabilities that are relevant to each species. In Nussbaum's view, apart from providing a solution to the problem of exclusion, this argument has an additional strength: it distinguishes between different species by appealing to the quantity and quality of each species' capabilities. Given that different capabilities are relevant to different species depending on each species' flourishing, the more complex a being is, the greater and more different the types of harm they can suffer (Nussbaum 2006, 361). If what we owe to others is affected by the quality and quantity of the harms and flourishing that they can experience, we arrive at the intuitively plausible implication that states and citizens owe more to

[70] Indeed, for example, Kimberly Smith criticizes Nussbaum on the grounds that her argument is overinclusive and cannot, therefore, be the subject of an overlapping consensus (Smith 2008, 205).
[71] According to Nussbaum's and Sen's development of the capability approach to social justice, capabilities are the real opportunities one has to be and do what it is reasonable for her to value.

humans than they do to donkeys, and more to donkeys than they do to flies. Simultaneously, however, by focusing on species and not on individuals within each species, we make differences in capacities and capabilities *within* a species irrelevant and are not, therefore, led to the exclusion of individual beings.

Although the main conclusions of this argument may fit with our intuitions, the argument itself is subject to a number of objections. First, there is an implausible implication: if the fact that each species can flourish in certain ways gives rise to enforceable duties to ensure that they do flourish, we reach the counterintuitive implication that states and citizens have enforceable duties towards all beings, including insects and other non-sentient animals. Even if these duties are weaker or narrower in scope than the duties we have towards humans or dolphins, it still seems both demanding and counterintuitive to suggest that there is an *enforceable* duty to ensure that all flies have access to the (minimal) capabilities of their species.

Nussbaum considers this objection and claims that in cases where there is no conflict of interests, the capabilities theorist should bite the bullet. For example, if a fly does not threaten human interests in any way, there is an enforceable duty to respect that fly's interest in living a decent life. Presumably, this not only implies that the fly should not be killed but also that it should be provided with opportunities to satisfy its basic needs. When there is a conflict of interests between humans and animals, however, we should give more weight to humans' interests. This is because, although there is an 'overlapping consensus on a basic minimum of capability for animals', there is no overlapping consensus on 'the equal dignity of creatures across species' (Nussbaum 2006, 383). That is, there is no overlapping consensus on whether humans and flies, for example, are of equal dignity. For

that reason, Nussbaum claims, we can treat the question of equal dignity as a metaphysical question on which there is reasonable disagreement.

However, we have already seen that we cannot use the idea of an overlapping consensus in order to justify the inclusion of animals. Even if there is overlapping consensus on the idea of dignity, that does not necessarily make that idea political rather than comprehensive. More importantly, as we have seen, it is implausible to suggest that there is overlapping consensus on the existence of *enforceable* duties to animals and even if there were, that would be irrelevant. For whether there is *actual* overlapping consensus cannot determine whether a group has rights or not. Therefore, the first objection to the capabilities argument – namely, that it leads to the implausible implication that there are enforceable duties to ensure each species' flourishing – retains its force.

Moreover, although Nussbaum presents this solution as an argument that can be given *within* the context of political liberalism, her argument relies on perfectionist judgments belonging to different comprehensive doctrines – it appeals to flourishing and therefore to a specific conception of the good.[72] By endorsing a specific account of flourishing for each species, states would implicitly pass a judgment on how humans (and members of other species) live their lives. For instance, the account seems to imply that a human should not want to live a life that resembles that of a chimpanzee more than that of a human – as both lives are standardly conceived by the capabilities theory. The reason why (most) capabilities matter is that they happen to be primary goods, i.e. goods that can promote all reasonable conceptions of the good – not that they are a crucial part of flourishing. It

[72] This is not just important for Rawlsian contractualism but for Nussbaum's own theory of justice as well, given that her theory uses the idea of public reason. This gives rise to an internal inconsistency in Nussbaum's theory, as the idea of public reason cannot support basing public policy on an account of flourishing.

follows that Nussbaum's argument cannot be used to achieve the inclusion of animals in Rawlsian contractualism.

6.2.6 An Intuitive Argument for Trusteeship

Finally, it has been suggested that the problem of exclusion can be solved if Rawlsian contractualism follows the trusteeship model, according to which the interests of those who lack the two moral powers are represented by trustees. For instance, even if the parties in the Original Position represent reasonable and rational persons, they should also act as trustees for those who lack the two moral powers.

Those who advance this argument point out that it is only by endorsing the trusteeship model that the theory can include children, infants, and humans with severe cognitive disabilities (Talbert 2006, 211). Indeed, my suggestion that the parties ought to take the interests of children and future generations into consideration can be viewed as an indirect endorsement of the trusteeship model. Yet if the theory appeals to trusteeship in order to include these groups, the argument goes, it should do so with regards to non-human animals as well; for the only way to exclude animals if the other excluded groups were represented through trustees would be by appealing to an arbitrary speciesist restriction (Talbert 2006, 209). Given that all of these groups lack the two moral powers, it is claimed that the only property that distinguishes animals from the rest is their species membership. And if all that matters for one's political status is one's capacity to participate in social cooperation, species membership is a morally irrelevant property. Thus, the argument concludes that if we endorse a trusteeship solution to the problem of other groups' – such as children's – exclusion, we must endorse a similar solution to the problem of animals' exclusion.

However, although I have indeed suggested that children, infants, unreasonable citizens, and citizens with severe cognitive disabilities are indeed to be included through trusteeship, the justification for their inclusion is not applicable to the case of animals. The inclusion of the former groups is warranted because they have the *potential* to develop the two moral powers and a related interest in fulfilling that potential, and because their parents and states have certain duties towards them that are defended through the idea of opacity respect. Yet neither of these two justifications for inclusion applies to the case of non-human animals. Therefore, it is not true that *if* we endorse the trusteeship model for some groups, we must endorse it for non-human animals too. Unless the proponents of this view advance a different justification for trusteeship for animals, the intuitive argument for trusteeship fails to warrant their inclusion.

6.3 An Argument from Compassion

Having examined a number of solutions that have been given to the problem of animals' exclusion, I now turn to defend a solution that is not, I think, vulnerable to the objections I have been levelling against other proposals. My argument is that once those included in the justificatory constituency (e.g. the parties in the Original Position) have settled questions about basic justice and constitutional essentials that concern those who can and do participate in social cooperation, they ought to face questions about how to treat those who cannot participate in social cooperation. For instance, the parties must deliberate on issues such as abortion, education, and criminal justice, which concern groups that lack the two moral powers – namely, children, foetuses, and unreasonable citizens. As we have seen, in response to questions about those who can develop the two moral powers, the parties ought to reflect on those beings' interest in developing the two moral powers and weigh it against the basic rights of those they represent (as they must do, for example, in the case of

abortion). Once they have addressed these questions, they must turn to the question of how they ought to treat those who will never develop the two moral powers – namely, individuals with severe cognitive disability and non-human animals. We have seen that the interests of the former group are taken into consideration through the idea of opacity respect. The interests of non-human animals, I will now suggest, would be taken into consideration by those included in the constituencies of justice and legitimacy because those included are reasonable, and reasonableness entails the capacity to feel and show compassion.

My argument here is twofold. First, there is a descriptive claim that, given our best understanding of normal human psychology, reasonable persons are expected to have the capacity to feel compassion. Recall that, according to Rawls, reasonableness tracks the capacity to understand the requirements of justice and the willingness to comply with these requirements. Although it is conceivable that, say, an agent of artificial intelligence could have this capacity without having the capacity to feel compassion, this does not seem possible once we focus on human beings. For instance, any sufficiently reasonable human being who complies with the principles of justice motivated purely by the right reasons (and not, for instance, out of fear of punishment) would be expected to have the capacity to put herself in other persons' shoes, and recognise their pleasure as good, their pain as bad, and their interests as sources of duties of compassion. There is no good reason to think that this person's capacity to do so would be restricted to the pleasure, pain, and broader basic interests of human beings but not of other sentient beings who have the capacity to similarly experience pleasure and pain. This implies, for instance, that reasonable persons would think that, given a sentient being's interest in not feeling pain, there is a *pro tanto* reason to refrain from causing pain to that being. Indeed, Rawls himself seems to endorse

this claim, albeit without justifying it, when he argues that 'the capacity for feelings of pleasure and pain and for the forms of life of which animals are capable clearly impose duties of compassion and humanity in their case' (Rawls 1971, 512). Although it is not clear why Rawls suggested that this capacity imposes *duties* of compassion and humanity, I suggest that it does lead us to the conclusion that, if my empirical claim is true, reasonable persons would be expected to respect and fulfil at least some of non-human animals' basic interests.

Apart from this *descriptive* claim that shows that the exclusion of non-human animals from the constituency of political persons would not be expected to have the extremely counterintuitive implications that we first saw, there is a second, *normative* claim that supports this. In fact, this second claim leads us to the normative conclusion that states *should* protect at least some of non-human animals' very basic negative interests. The claim in question is that, as we have seen in the chapter on children, given the link between reasonableness and compassion, states ought to cultivate compassion in their citizens, in order to enable them to become reasonable. And if compassion is cultivated as a political virtue in citizens, we would expect citizens to naturally feel compassion towards non-human sentient beings too. This line of argument can be traced back to Kant's account of indirect duties: even if there are no direct duties towards animals, there are indirect duties because behaving in certain ways towards animals predisposes us to behave in similar ways towards humans (Kant (1997 [1784–5]), 212; Nussbaum 2006, 329). Indeed, for instance, if one is cruel to animals and tortures them for fun, one is likely to (want to) behave cruelly towards humans as well (Stupperich and Strack, 2016; Sereny, 1972; MacDonald, 1963; Hellman and Blackman, 1966). Since public policy should seek to cultivate reasonableness and since traits such as compassion, empathy, and kindness have a positive influence on

human beings' reasonableness, we can conclude that states should foster these traits. And we can plausibly expect that disregarding animals' basic interests, as one does when one tortures an animal for fun, is not consistent with the development of the traits associated with reasonableness; for instance, when faced with a non-human animal in pain, a compassionate human being is naturally expected to feel compassion towards that animal. Moreover, it seems plausible that apart from merely cultivating reasonableness, states should *enforce* certain duties of compassion; for instance, acts such as torturing animals for fun should be banned, as they seem entirely inconsistent with the development of reasonableness. Therefore, we see that some of the interests of animals would be protected both because political liberal states ought to cultivate traits such as compassion *and* because states ought to ban certain acts that frustrate these interests, whenever permitting these acts allows room for the cultivation of tendencies associated with unreasonableness.

To be sure, duties of compassion are not the same as duties of justice. The former are less urgent than the latter and can, by extension, be outweighed by less weighty considerations than the latter can. Thus, I am not claiming that reasonable persons have equal duties to animals and humans or that they must satisfy all basic interests that animals may have; rather, the limited claim that my argument makes is that there is a weighty *pro tanto* reason for reasonable persons to take sentient animals' basic interests into account when assessing the justifiability of certain laws and policies. While this claim is limited, it shows that Rawlsian contractualism can avoid the problem of counterintuitive implications that is created by the total exclusion of animals from the realms of justice and legitimacy. And it should be noted that this claim does not preclude a greater range of legally enforceable duties towards animals; the question of further duties remains to be settled by procedural solutions such as democratic voting.

Nussbaum has objected to compassion-based arguments by arguing that the emotion of compassion excludes the notions of blame and responsibility. For example, a person's distress arouses compassion in the same way regardless of whether she is the victim of a crime or the victim of bad brute luck for which no one is responsible (Nussbaum 2006, 336). Thus, according to Nussbaum, the appeal to compassion would have certain counterintuitive implications. Imagine, for instance, that person A harms person B; this is bad not simply because person B's well-being has diminished but because they have been wronged. For that reason, person A ought to be held responsible for that harm. As a result of focusing solely on compassion when contemplating on the reasons there are to alleviate B's suffering, Nussbaum argues, this dimension of responsibility is lost. That is, we would be treating a case of interpersonal harm as a case of impersonal harm. In doing so, we would fail to view individuals as agents who have duties towards others and who can be blamed when failing to fulfil these duties. If Nussbaum's objection to compassion-based arguments were plausible, then compassion would not be sufficient for the full inclusion of animals because it would exclude the ideas of blame and responsibility.

It is difficult to see the force of this objection, however. For we may plausibly think that if we have duties of compassion, we have duties to prevent suffering. There is no reason to think that compassion applies only to cases where suffering has already taken place; it might also give rise to preventive duties, given that we ought to feel compassion when we imagine someone's suffering. If compassion *can* give rise to (both preventive and reactive) duties, it follows that individuals are responsible for fulfilling these duties and that they ought to be blamed when they violate their duties. Thus, the idea of compassion does not necessarily exclude the notions of blame and responsibility.

We thus reach the conclusion that those included in the justificatory constituency would and should take animals' interests into consideration in their deliberations. Despite precluding an appeal to comprehensive doctrines, Rawlsian contractualism can still be concerned with the kinds of citizens that all individuals ought to be; torturing non-members for fun is inconsistent with an attitude of respect for beings who have interests. I now turn to the question of what taking animals' interests into consideration in the way suggested by these arguments actually implies.

Given that any basic interests that animals have might often clash with the interests of those who count as full persons, it may seem that the theory would have to require the sacrifice of animals' interests in many cases. However, although it is not possible to provide a full account of the range of duties that arise – partly because their justification would have to be contextual, in the sense that it would depend on the interests at stake in each case – we should notice that taking someone's interests seriously means that we weigh those interests against competing interests and assess their significance. For instance, the basic interest that animals have in living may not give rise to an absolute right to life, yet it ought to be weighed against competing interests, rather than merely ignored. Taking that interest seriously means that the competing interest that one may have in eating animals if there are sufficiently good alternatives cannot be viewed as stronger. Thus, although it may be disappointing that animals cannot be fully included in a Rawlsian theory of justice and legitimacy through a rights-based approach, the theory cannot be accused of excluding them completely or of having reprehensible implications, such as allowing torture for fun.

It should be noted, however, that the duties that are generated towards animals from society's interest in being populated by reasonable citizens are broader and stronger in the case of pets. Although society has an interest in respecting the basic interests of animals, the duties that arise from this interest should not be distributed equally in the case of pets. That is, pet owners should have a greater share of duties and these duties should be more stringent than they would be in the case of wild animals.

To see why, notice that, much as in the case of parents, pet owners have special duties towards their own pets because they bear moral responsibility for their needs and, in some cases, for their existence. For instance, by domesticating an animal, humans both prevent animals from developing the capacities that they need in order to cover their needs *and* become responsible for the development of new needs, such as the need for human companionship and love. And, as we have seen, political liberal states must hold individuals responsible for the ends that they have and for the choices that they make on the basis of those ends. Therefore, the moral responsibility that pet owners have towards their pets implies that they have a greater range of duties towards them than they would if these pets had never been domesticated.

At this point, it might be objected that we may not draw this conclusion from the analogy between children and pets because it would not be permissible for states to *force* pet owners to comply with a great range of duties. For, as we have seen, given that children are future persons, the duties that exist towards them are duties of justice, while duties towards animals are less weighty duties of compassion. I suggest, however, that this objection misfires. To see why, suppose that a pet is abandoned and is not capable of satisfying its own interests in the ways that a wild animal would. This animal's suffering will give rise

to duties of compassion that would fall on the rest of society. If a state remained indifferent towards that suffering, it would essentially fail to promote the civic attitudes that are required for and by reasonableness, which include compassion. On the other hand, however, neither would it be fair for states to merely let the rest of society bear the duties of compassion that have been created by the abandonment of that pet. For fairness requires that states recognise the moral responsibility that pet owners bear, in the sense that they have created the additional needs this animal may have and they have also caused this animal's inability to satisfy its own needs. So, the justification for the claim that a greater range of duties towards pets should be enforced is that a failure to fulfil these duties on the part of those who are morally responsible for them would be socially harmful. Thus, we can conclude that the duties that arise from animals' interests in the case of pets are broader, more demanding, and distributed differently from the duties of compassion that arise in the case of wild animals.

6.4 Conclusion

In this chapter, I have argued that non-human animals seem to be excluded from Rawlsian contractualism because they lack both the two moral powers and the potential to develop them in the future. Their exclusion implies that citizens and states in Rawlsian contractualism have absolutely no duties with regard to them, which further implies that even acts such as torturing animals for fun cannot be deemed impermissible. Given that this implication is intuitively viewed as morally abhorrent, the exclusion of animals leaves Rawlsian contractualism vulnerable to the problem of counterintuitive implications.

In the first section of the chapter, I examined and rejected a number of potential solutions that have been offered in response to the question of animals' exclusion. These solutions

were: the intuitive equality argument; the claim that we have relationships with animals for which the two moral powers are not required; the argument that contractualism could include an appeal to altruism; the overlapping consensus argument; the capabilities and species membership argument; and the intuitive argument in favour of trusteeship. After explaining why these suggestions cannot, in fact, resolve the problem of animals' exclusion, I offered an alternative solution. More specifically, in the second part of the chapter, I argued that there are duties of compassion that establish a *pro tanto* reason to take animals' basic interests into consideration in the justificatory constituency's deliberation. These duties exist because, given our understanding of human psychology, the cultivation of citizens' reasonableness requires the cultivation of a concern for all sentient beings' basic interests. Although these duties of compassion are less weighty than the duties that I have defended towards other groups that may seem to be excluded from Rawlsian contractualism, they nonetheless address the question of exclusion in a way that prevents the theory from having extremely counterintuitive implications.

Chapter 7

Conclusion

All theories of justice and legitimacy face the question of who is included in society and under what terms. The questions that arise include who counts as a free and equal citizen, who has full participation and representation rights in politics, and what kinds of enforceable rights and duties different groups and individuals have. It has been a key assertion of this thesis that social contract approaches to justice and legitimacy face a particularly pressing challenge when it comes to answering the question of inclusion. For, intuitively, the exclusion of some individuals and groups seems to be a corollary of the social contractualist approach. This is because this approach ascribes justice or legitimacy to a social arrangement insofar as it can be regarded as the product of the (actual – expressed or tacit – or hypothetical) consent of specific parties, thereby excluding all those whose consent is not required in order to view an arrangement as legitimate or just.

This thesis has focused on a specific account of the social contractualist approach – Rawlsian contractualism – and on the ways in which the question of exclusion hallenges that approach. Following the social contract tradition, Rawlsian contractualism views society as social cooperation with the purpose of mutual advantage. For that reason, those who are included on equal terms in the theory are those who have the capacity to contribute to social cooperation. In order to have that capacity, according to Rawlsian contractualism, one must be sufficiently reasonable and rational. They must be reasonable in the sense that they understand the requirements of justice and in that they are willing to comply with

these requirements. And they must be rational in the sense that they have the capacity to form, revise, and pursue a conception of the good. These two requirements ground the political conception of personhood. That is, in order to count as a person in the political sense, one must be sufficiently reasonable and rational.

Moreover, given the constraints set by public reason, the theory cannot rely on an alternative, metaphysical conception of personhood in order to view those who lack these capacities as persons. Because metaphysical conceptions of personhood are parts of comprehensive doctrines, any claims about such conceptions would not be publicly justifiable. And given Rawlsian contractualism's commitment to public justifiability, the theory cannot appeal to such conceptions of personhood in order to justify the inclusion of those who lack the two moral powers in society as full persons.

This gives rise to two categories of exclusion. First, there are those who will develop the two moral powers sufficiently only if those who count as persons in the political sense act or refrain from acting in certain ways. I explored this category of exclusion in the first part of the thesis, focusing on children, foetuses, and unreasonable persons. Second, there are those who will *never* develop the two moral powers to the requisite degree, such as individuals with severe cognitive disabilities and non-human animals. This case of exclusion was explored in the second part of the thesis.

I have argued throughout the thesis that there are two broad, related questions of exclusion that arise with regards to all groups that lack the two moral powers. The first question concerns their exclusion from the constituency of legitimacy. Because the liberal principle of legitimacy requires that political power be justifiable to reasonable and rational persons,

it follows that political power need not be justifiable to those who lack the two moral powers. Relatedly, the second question concerns these groups' exclusion from the constituency of justice. This question can be interpreted in two ways. First, there is a question about who is represented in the constituency that decides on the principles of justice. In Justice as Fairness, it translates into the question of who is represented in the Original Position. As we have seen, this question follows naturally from the question of exclusion from the constituency of legitimacy. Given that only reasonable and rational persons are represented in the Original Position, it follows that the interests of those who lack the two moral powers may not be taken into consideration in the parties' deliberation. The second interpretation of the question of exclusion from the constituency of justice is about those to whom the principles of justice apply. If the principles are meant to distribute the benefits and burdens of social cooperation *among those who have the capacity to cooperate*, they only apply to those who have this capacity – namely, to reasonable and rational persons. Therefore, the principles of justice as they stand, in this case the principles of Justice as Fairness, may not apply to those who do not possess the two moral powers.

I have argued that, unless Rawlsians address these questions of exclusion, the theory faces a number of problems. To begin with, the theory faces the problem of internal inconsistency. This is because the exclusion of some individuals who *currently* lack the two moral powers but who have the *potential* to develop them in the future reveals temporal partiality. But given that the theory is otherwise committed to the idea of temporal impartiality, the implicit appeal to temporal partiality reveals an internal inconsistency. And if the idea of temporal partiality is independently objectionable, this implicit appeal also makes the theory vulnerable to the external objection that it is an implausible theory. Moreover, the question of exclusion makes the theory vulnerable to the charge of

indeterminacy – that is, the charge that the theory cannot reach *any* conclusions on practical questions regarding those who are excluded. The thought here is that the theory can *only* formulate arguments that appeal to public reasons and that reaching a conclusion on any issue regarding those excluded would require to implicitly appeal to non-public reasons. Lastly, the question of exclusion gives rise to the charge of counterintuitive implications. If Rawlsian contractualism excludes the interests of those excluded from the constituencies of justice and legitimacy, it may have extremely objectionable implications. This seems problematic both in itself, and because the theory is committed to the idea of reflective equilibrium and to therefore having implications that are in line with our considered judgements. Thus, we see that unless Rawlsians respond to the question of exclusion, Rawlsian contractualism faces a number of pressing objections; indeed, we have seen that a number of theorists have pressed these objections in one form or another, albeit typically not as problems of *exclusion*.

The main purpose of this thesis has been to show *how* Rawlsians can respond to this question. In doing so, I have argued that the question of all groups' exclusion cannot be addressed by a single solution and have dealt with each case on an individual basis, showing how Rawlsian contractualism can give plausible policy-prescriptions on issues pertaining to each group. Taking Rawls's solution to intergenerational justice as a starting point, I have argued that the main way of avoiding the exclusion of those who have the potential to become reasonable and rational in the future if current persons act or refrain from acting in certain ways is by appealing to the interest in becoming reasonable and rational. I have argued that this interest is *politically* justified and can therefore be appealed to without violating the requirements of public reason. I have also suggested that, given Rawlsian contractualism's commitment to temporal impartiality, those included in the

constituencies of legitimacy and justice ought to take this interest into consideration in their deliberations. This externally imposed constraint on them ensures that there are no objectionable instances of exclusion. In this way, children, removed, viable foetuses, and unreasonable persons can be included in the theory. The policy implications of my argument are often controversial. Among others, for instance, we have seen that parents should be required to take parenting lessons that will teach them the range of techniques that are required as part of their duty to satisfy their children's interest in developing the two moral powers. Similarly, we have seen that, with regards to unreasonable persons, this argument implies that paternalistically justified policies such as mandatory education or neurointerventions, which would otherwise be presumptively wrong, are sometimes required in order to satisfy these persons' interest in developing their reasonableness. Therefore, although it is true that these excluded groups are not included under the same terms that fully reasonable and rational persons are, their basic interests are not excluded from the considerations that those included in the justificatory constituency (and, by extension, in the Original Position) appeal to in their deliberations.

Apart from this interest-based argument, which only applies to those who have the potential to become reasonable and rational in the future, I have argued that there are a number of other publicly justifiable reasons that are relevant to each case and which can help Rawlsians respond to each case of exclusion. These reasons are different in the case of individuals with cognitive disabilities than that of non-human animals. In the former case, I have defended a twofold argument. First, I have argued that, because it is often impossible to determine whether individuals with cognitive disabilities have the potential to develop the two moral powers, states should err on the side of caution. Second, I have argued that, even if there is no such potential, creating children gives rise to parental, state, and citizens'

duties towards children. Children with cognitive disabilities, I have claimed, could not be viewed as an exception to the duty-based argument, because those included in the justificatory constituency ought to show opacity respect to all children when determining the group to whom these duties are owed. Lastly, I have examined the question of exclusion as it applies to non-human animals. I have suggested that there are duties of compassion that ground a *pro tanto* reason for those justificatory constituency to take animals' basic interests into account. These duties are justified within the context of public reason because, I have claimed, the cultivation of citizens' reasonableness requires the cultivation of a concern for all sentient beings' basic interests. I hope to have shown that these arguments collectively respond to the question of exclusion as that applies to each particular case, thereby strengthening the case for Rawlsian contractualism.

Bibliography

Abbey, R. (2007) 'Rawlsian Resources for Animal Ethics', *Ethics and the Environment*, 12(1), pp. 1—22.

Ainsworth, M. D. S., Bell, S. M., and Stayton, D. J. (1972) 'Individual differences in the development of some attachment behaviors', *Merrill-Palmer Quarterly*, 18 (2), pp. 123—143.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., and Wahl, S. (1978) *Patterns of Attachment*. Hillsdale, NJ: Erlbaum.

Allinsmith, W. and Greening, T. C. (1955) 'Guilt over Anger as Predicted from Parental Discipline: A Study of Superego Development', *American Psychologist*, 10, p. 320.

Anderson, E. (1999) 'What Is the Point of Equality?', *Ethics*, 109 (2), pp. 287—337.

Andersson, E. (2011) 'Political Liberalism and the Interests of Children: A Reply to Timothy Michael Fowler', *Res Publica*, 17 (3), pp. 291—296.

Arneson, R. (1989) 'Equality and Equal Opportunity for Welfare', *Philosophical Studies*, 56 (1), pp. 77—93.

Badano, G. (2014) 'Political Liberalism and the Justice Claims of the Disabled: A Reconciliation', *Critical Review of International Social and Political Philosophy*, 17 (4), pp. 401—422.

Barn, G. (2019) 'Can Medical Interventions Serve as "Criminal Rehabilitation"?', *Neuroethics*, 12, pp. 85—96.

Batson, D. C. (1991) *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale, NJ: Lawrence Erlbaum.

Batson, D. C. (2009) 'These Things Called Empathy', in J. Decety and W. Ickes (eds.), *The Social Neuroscience of Empathy*. Cambridge, MA: MIT Press.

Becker, L. C. (2005) 'Reciprocity, Justice, and Disability', *Ethics*, 116 (1), pp. 9—39.

Berkowitz, M. W. and Grych, J. A. (1998) 'Fostering Goodness: Teaching Parents to Facilitate Children's Moral Development', *Journal of Moral Education*, 27 (3), pp. 371—391.

Bernstein, M. (1997) 'Contractualism and Animals', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 86(1), pp. 49—72.

Bérubé, M. (2010) ‘Equality, Freedom, and/or Justice for All: A Response to Martha Nussbaum’, in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Bowlby, J. (1980) *Loss, Sadness and Depression: vol. 3, Attachment and Loss*. New York: Basic Books.

Brennan, J. (2016) *Against Democracy*. Princeton, NJ: Princeton University Press.

Brighouse, H. (1998) ‘Civic Education and Liberal Legitimacy’, *Ethics*, 108 (4), pp. 719—745.

Brighouse, H. and Swift, A. (2014) *Family Values: The Ethics of Parent-Child Relationships*. Princeton, NJ: Princeton University Press.

Bublitz, J. C. and Merkel, R. (2014) ‘Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination’, *Criminal Law and Philosophy*, 8, pp. 51—77.

Carruthers, P. (1992) *The Animals Issue: Moral Theory in Practice*. Cambridge: Cambridge University Press.

Carter, I. (2011) ‘Respect and the Basis of Equality’, *Ethics*, 121 (3), pp. 538—571.

Carter, I. (2013) ‘Basic Equality and the Site of Egalitarian Justice’, *Economics & Philosophy*, 29 (1), pp. 21—41.

Casal, P. (2003) ‘Is Multiculturalism Bad for Animals?’ *Journal of Political Philosophy*, 11(1), pp. 1—22.

Casal, P. and Williams, A. (1995) ‘Rights, Equality and Procreation’, *Analyse & Kritik*, 17 (1), pp. 93—116.

Chew, C., Douglas, T., and Faber, N. (2018) ‘Biological Interventions for Crime Prevention’, in D. Birks and T. Douglas (eds.), *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford: Oxford University Press.

Clarke-Stewart, K. A. (1973) ‘Interactions Between Mothers and their Young Children’, *Monographs of the Society for Research in Child Development*, 38 (6/7), pp. 1—109.

Clayton, M. (2000) ‘The Resources of Liberal Equality’, *Imprints*, 5, pp. 63—84.

Clayton, M. (2006) *Justice and Legitimacy in Upbringing*. Oxford: Oxford University Press.

Clayton, M. and Chan, T. W. (2006) ‘Should the Voting Age Be Lowered to Sixteen? Normative and Empirical Considerations’, *Political Studies*, 54 (10), pp. 533—558.

Cleare, A. J. and Bond, A. J. (1995) ‘The Effect of Tryptophan Depletion and Enhancement on Subjective and Behavioural Aggression in Normal Male Subjects’, *Psychopharmacology*, 118, pp. 72-–81.

Cohen, G. A. (1989) ‘On the Currency of Egalitarian Justice’, *Ethics*, 99 (4), pp. 906—944.

Crockett, M. J., Apergis-Schoute, A., Herrmann, B., Lieberman, M. D., Müller, U., Robbins, T. W., and Clark, L. (2013) ‘Serotonin Modulates Striatal Responses to Fairness and Retaliation in Humans’, *The Journal of Neuroscience*, 33, pp. 3505–-3513.

Dahl, N. O. (1974) ‘’Ought’ Implies ‘Can’ and Deontic Logic’, *Philosophia*, 4, pp. 485–-511.

Dougherty, D. M., Bjork, J. M., Marsh, D. M., and Moeller, F.G. (1999) ‘Influence of Trait Hostility on Tryptophan Depletion-Induced Laboratory Aggression’, *Psychiatry Research*, 88, pp. 227-–32.

Dworkin, R. (1994) *Life’s Dominion*. New York: Vintage Books.

Dworkin, R. (2000) *Sovereign Virtue*. Cambridge, MA.: Harvard University Press.

Egeland, B. and Sroufe, L. A. (1981) ‘Attachment and Early Maltreatment’, *Child Development*, 52 (1), pp. 44—52.

Eisenberg, N. and Mussen, P. H. (1989) *The Roots of Prosocial Behavior in Children*. New York: Cambridge University Press.

Elliot, R. (1984) ‘Rawlsian Justice and Non-Human Animals’, *Journal of Applied Philosophy*, 11 (1), pp. 95—106.

Erikson, E. H. (1950) *Children and Society*. New York: Norton.

Evers, W. M. (1979) ‘Rawls and Children’, *Journal of Libertarian Studies*, 2 (2), pp. 109—114.

Folbre, N. (1994) ‘Children as Public Goods’, *The American Economic Review*, 84 (2), pp. 86—90.

Folbre, N. (2001) *The Invisible Heart*. New York: New Press.

Fowler, T. (2010) ‘The Problems of Liberal Neutrality in Upbringing’, *Res Publica*, 16 (4), pp. 367—381.

Fowler, T. (2014a) ‘The Status of Child Citizens’, *Politics, Philosophy & Economics*, 13 (1), pp. 93—113.

Fowler, T. (2014b) ‘Perfectionism for Children, Anti-Perfectionism for Adults’, *Canadian Journal of Philosophy*, 44 (3-4), pp. 305—323.

Frankfurt, H. G. (1971) ‘Freedom of the Will and the Concept of a Person’, *Journal of Philosophy*, 68 (1), pp. 5—20.

Freeman, S. (2002) ‘Congruence and the Good of Justice’, in S. Freeman (ed.), *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press.

Freeman, S. (2007) *Justice and the Social Contract: Essays in Rawlsian Political Philosophy*. New York: Oxford University Press.

Friberg-Fernros, H. (2010) ‘Abortion and the Limits of Political Liberalism’, *Public Reason*, 2 (1), pp. 27—42.

Friedman, M. (2000) ‘John Rawls and the Political Coercion of Unreasonable People’ in Victoria Davion and Clark Wolf (eds.), *The Idea of a Political Liberalism*. Oxford: Rowman and Littlefield.

Garner, R. (2003a) ‘Animals, Politics and Justice: Rawlsian Liberalism and the Plight of Non-Humans’, *Environmental Politics*, 12(2), pp. 3—22.

Garner, R. (2003b) ‘Political Ideologies and the Moral Status of Animals’, *Journal of Political Ideologies*, 8(2), pp. 233—246.

Gaus, G. (1996) *Justificatory Liberalism*. Oxford: Oxford University Press.

Gaus, G. (2009) ‘The Moral Foundations of Liberal Equality’, in Thomas Christiano and John Christman (eds.) *Debates in Contemporary Political Philosophy*. Oxford: Blackwell.

Gaus, G. (2011) *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press.

Gibbs, J.C. (1987) ‘Social Processes in Delinquency: The Need to Facilitate Empathy as well as Sociomoral Reasoning’, in W.M. Kurtines and J.L. Gewirtz (eds.) *Moral Development Through Social Interaction*. New York: Wiley.

Goodin, R. (2005) ‘Responsibilities for Children’s Well-Being’, M. Prior and S. Richardson (eds.), *No Time to Lose: The Wellbeing of Australia’s Children*. Melbourne: Melbourne University Press.

Gordon, H. and Grubin, D. (2004) ‘Psychiatric Aspects of the Assessment and Treatment of Sex Offenders’, *Advances in Psychiatric Treatment*, 10 (1), pp. 73—80.

Green, L. (1995) ‘Internal Minorities and Their Rights’, in W. Kymlicka (ed.), *The Rights of Minority Cultures*. Oxford: Blackwell.

Harter, S. (1997) ‘The Development of Self-Representations, in N. Eisenberg (ed.), *Handbook of Child Psychology*, 5th edn, vol. III. New York: Wiley.

Hartley, C. (2009) ‘Justice for the Disabled: A Contractualist Approach’, *Journal of Social Philosophy*, 40 (1), pp. 17—36.

Hayward, T. (2013) ‘On Prepositional Duties’, *Ethics*, 123(2), pp. 264—291.

Hellman, D. S. and Blackman, N. (1966) ‘Enuresis, Firesetting and Cruelty to Animals: A Triad Predictive of Adult Crime’, *The American Journal of Psychiatry*, 122(22), pp. 1431—1435.

Hoffman, M. L. (1983) ‘Affective and Cognitive Processes in Moral Internalization’, in E.T. Higgins, D.N. Ruble, and W.W. Hartup (eds.), *Social Cognition and Social Development: A Sociocultural Perspective*. New York: Cambridge University Press.

Hoffman, M.L. and Saltzstein, H.D. (1967) ‘Parent Discipline and the Child's Moral Development’, *Journal of Personality and Social Psychology*, 5 (1), pp. 45—57.

Howard, J. (2017) ‘Punishment as Moral Fortification’, *Law and Philosophy*, 36 (1), pp. 45—75.

Hume, D. (2000 [1738]) *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton. New York: Oxford University Press.

Jaworska, A. (2010) ‘Caring and Full Moral Status’, in E. F. Kittay and L. Carlson, (eds.) *Cognitive Disability and its Challenge to Moral Philosophy* (Oxford: Wiley-Blackwell).

Kant, I. (1998 [1793]) *Religion within the Boundaries of Mere Reason*, translated by Allen Wood and George Di Giovanni. Cambridge: Cambridge University Press.

Kant, I. (1997 [1784–5]) ‘Moral Philosophy: Collin’s Lecture Notes’, in *Lectures on Ethics*, (Cambridge Edition of the Works of Immanuel Kant), P. Heath and J.B. Schneewind (eds. and trans.). Cambridge: Cambridge University Press, pp. 37–-222.

Kant, I. (1971 [1797]) *The Doctrine of Virtue: Part II of the Metaphysics of Morals*, translated by Mary J. Gregor. Philadelphia, PA: University of Pennsylvania Press.

Kirby, N. (2018) ‘Two Concepts of Equality’, *Res Publica*, 24 (3), pp. 297—318.

Kittay, E. F. (2010) ‘The Personal Is Philosophical Is Political: A Philosopher and Mother of a Cognitively Disabled Person Sends Notes from the Battlefield’, in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Klosko, G. (2000) *Democratic Procedures and Liberal Consensus*. Oxford: Oxford University Press.

Kochanska, G., Murray, K. and Coy, K. C. (1997) ‘Inhibitory Control as a Contributor to Conscience in Childhood: From Toddler to School Age’, *Child Development*, 68 (2), pp. 263—277.

Kymlicka, W. (1990) ‘Two Theories of Justice’, *Inquiry*, 33(1), pp. 99—119.

Liao, M. S. (2015) *The Right to Be Loved*. Oxford: Oxford University Press.

MacDonald, J. M. (1963) 'The Threat to Kill', *American Journal of Psychiatry*, 120(2), pp. 124—130.

Main, M. Tomasini, L., and Tolan, W. (1979) 'Differences Among Mothers of Infants Judged to Differ in Security', *Developmental Psychology*, 15 (4), pp. 472—3.

Maletzky, B. M. and Field, G. (2003) 'The Biological Treatment of Dangerous Sexual Offenders: A Review and Preliminary Report of the Oregon Pilot Depo-Provera Program', *Aggression and Violent Behavior*, 8 (4), pp. 391–-412.

Mangham, L. J., Petrou, S., Doyle, L. W., Draper E. S., and Marlow, N. (2009) 'The Cost of Preterm Birth Throughout Childhood in England and Wales, *Pediatrics*, 123 (2), e312—27.

McMahan, J. (2002) *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.

McMahan, J. (2006) 'Paradoxes of Abortion and Prenatal Injury', *Ethics*, 116 (4), pp. 625—655.

McMahan, J. (2010) 'Cognitive Disability and Cognitive Enhancement', in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Mill, J. S. (1979 [1861]) *Utilitarianism*. Indianapolis: Hackett Publishing.

Moeller, F. G., Dougherty, D. M., Swann, A. C., Collins, D., Davis, C. M. and Cherek, D. R. (1996) 'Tryptophan Depletion and Aggressive Responding in Healthy Males', *Psychopharmacology*, 126, pp. 97–-103.

Munoz-Dardé, V. and Sinclair, T. (forthcoming) *The Routledge Guidebook to Rawls's A Theory of Justice*.

Nathan, C. (2011) 'Need There Be a Defence of Equality?', *Res Publica*, 17, pp. 211—225.

Nussbaum, M. (1995) 'Human Capabilities, Female Human Beings', in M. C. Nussbaum and J. Glover (eds.), *Women, Culture, and Development: A Study of Human Capabilities*. Oxford: Oxford University Press.

Nussbaum, M. (2006) *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, MA: Harvard University Press.

Nussbaum, M. (2007) 'Radical Evil in Liberal Democracies: The Neglect of the Political Emotions', in T. Banchoff (ed.), *Democracy and the New Religious Pluralism*. New York: Oxford University Press.

Nussbaum, M. (2010) 'The Capabilities of People with Cognitive Disabilities', in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Nussbaum, M. (2011) 'Perfectionist Liberalism and Political Liberalism', *Philosophy & Public Affairs*, 39, pp. 3—45.

Nussbaum, M. (2013) *Political Emotions: Why Love Matters for Justice*. Cambridge, MA: Belknap Press of Harvard University Press.

Olsaretti, S. (2013) 'Children as Public Goods?', *Philosophy and Public Affairs*, 41 (3), pp. 226—258.

Park, K. A. and Waters, E. (1989) 'Security of Attachment and Preschool Friendships', *Child Development*, 60 (5), pp. 1076—1081.

Parke, R. D. and Walters, R. H. (1967) 'Some Factors Determining the Efficacy of Punishment for Inducing Response Inhibition', *Monographs of the Society for Research in Child Development*, 32 (1), pp. 1—45.

Phillips, D. (1998) 'Contractualism and Moral Status', *Social Theory and Practice*, 24(2), pp. 183—204.

Plunkett, D. (2016) 'Justice, Non-Human Animals, and the Methodology of Political Philosophy', *Jurisprudence*, 7(1), pp. 1—29.

Prabhakar, D. and Balon, R. (2010) 'How Do SSRIs Cause Sexual Dysfunction?', *Current Psychiatry*, 9 (12), pp. 30–-34.

Pringle, M. K. (1986) *The Needs of Children*. Abington: Routledge.

Quong, J. (2011) *Liberalism without Perfection*. Oxford: Oxford University Press.

Rawls, J. (1971) *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rawls, J. (1999) *The Law of Peoples*. Cambridge, MA: Harvard University Press.

Rawls, J. (2001) *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press.

Rawls, J. (2005) *Political Liberalism: Expanded Edition*. New York: Columbia University Press.

Regan, T. (1988) *The Case for Animal Rights*. London: Routledge.

Reidy, D. A. (2000) 'Rawls's Wide View of Public Reason: Not Wide Enough', *Res Publica*, 6, pp. 49—72.

Rodriguez, L. M., DiBello, A. M., Øverup, C. S., and Neighbors, C. (2015) 'The Price of Distrust: Trust, Anxious Attachment, Jealousy, and Partner Abuse', *Partner Abuse*, 6 (3), pp. 298—319.

Ross, W. D. (1930) *The Right and the Good*. Oxford: Oxford University Press.

Rowlands, M. (1997) ‘Contractarianism and Animal Rights’, *Journal of Applied Philosophy*, 14 (3), pp. 235—247.

Rowlands, M. (2013) ‘Animal Rights’, in Hugh LaFollette (ed.), *The International Encyclopedia of Ethics*. Oxford: Blackwell Publishing.

Sandel, M. (2005) *Public Philosophy: Essays on Morality in Politics*. Cambridge: Harvard University Press.

Satz, D. (2007) ‘Equality, Adequacy, and Education for Citizenship’, *Ethics*, 117 (4), pp. 623—648.

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

Scanlon, T. M. (2004) ‘Adjusting Rights and Balancing Values’, *Fordham Law Review*, 72 (5), p. 1481.

Schapiro, T. (1999) ‘What Is a Child?’, *Ethics*, 109 (4), pp. 715—738.

Schwartzman, M. (2004) ‘The Completeness of Public Reason’, *Politics, Philosophy & Economics*, 3, pp. 191—220.

Sereny, G. (1972) *The Case of Mary Bell – A Portrait of a Child Who Murdered*. London: Pimlico.

Shakespeare, T. (2017) ‘The Social Model of Disability’, in L. J. Davis (ed.), *The Disability Studies Reader*, 5th Edition. Abingdon: Routledge.

Shaw, D. M. (2011) ‘Justice and the Fetus: Rawls, Children, and Abortion’, *Cambridge Quarterly of Healthcare Ethics*, 20, pp. 93—101.

Silvers, A. and Francis, L. P. (2005) ‘Justice through Trust: Disability and the ‘Outlier Problem’ in Social Contract Theory’, *Ethics*, 116, pp. 40—76.

Silvers, A. and Francis, L. P. (2009) ‘Thinking about the Good: Reconfiguring Liberal Metaphysics (or Not) for People with Cognitive Disabilities’, *Metaphilosophy*, 40 (3-4), pp. 475—498.

Singer, P. (2010) ‘Speciesism and Moral Status’, in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Sinnott-Armstrong, W. (1984) “Ought’ Conversationally Implies ‘Can’’, *The Philosophical Review*, 93, pp. 249–-261.

Smith, K. K. (2008) ‘Animals and the Social Contract’, *Environmental Ethics*, 30(2), pp. 195—207.

Stark, C. A. (2007) ‘How to Include the Severely Disabled in a Contractarian Theory of Justice’, *Journal of Political Philosophy*, 15, pp. 127–-45.

Stark, C. A. (2010) 'Respecting Human Dignity: Contract versus Capabilities', in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Stupperich, A. and Strack, M. (2016) 'Among a German Sample of Forensic Patients, Previous Animal Abuse Mediates Between Psychopathy and Sadistic Actions', *Journal of Forensic Sciences*, 61 (3), pp. 699—705.

Swanson, J. (2011) 'Contractualism and the Moral Status of Animals', *Between the Species*, 14(1), pp. 1—17.

Talbert, M. (2006) 'Contractualism and Our Duties to Nonhuman Animals', *Environmental Ethics*, 28(2), pp. 201—215.

Theofilopoulou, A. (2015) 'Does Luck Egalitarianism Lose its Appeal in the Face of Genetic Engineering?', *Bioethica*, 1 (2), pp. 11—23.

Theofilopoulou, A. (2019) 'Punishment as Moral Fortification and Non-Consensual Neurointerventions', *Law and Philosophy*, 38 (2), pp. 149—167.

Thomson, J. J. (1971) 'A Defense of Abortion', *Philosophy & Public Affairs*, 1, pp. 47—66.

Tucker, C. and MacDonald, C. (2004) 'Beastly Contractarianism? A Contractarian Analysis of the Possibility of Animal Rights', *Essays in Philosophy*, 5(2), article 31.

Van DeVeer, D. (1979) 'Of Beasts, Persons and the Original Position', *The Monist*, 62(3), pp. 368—377.

Walker, L. J. and Taylor, J. H. (1991) 'Family Interactions and the Development of Moral Reasoning', *Child Development*, 62 (2), pp. 264—283.

Weiner, I. B. (1980) 'Psychopathology in Adolescence', in J. Adelson (ed.), *Handbook of Adolescent Psychology*. New York: Wiley.

Weithman, P. (2011) *Why Political Liberalism? On John Rawls's Political Turn*. Oxford: Oxford University Press.

Wenar, L. (1995) 'Political Liberalism: An Internal Critique', *Ethics*, 106 (1), pp. 32—62.

Wikler, D. (2010) 'Cognitive Disability, Paternalism, and the Global Burden of Disease', in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Williams, A. (2000) 'The Alleged Incompleteness of Public Reason', *Res Publica*, 6, pp. 199—211.

Williams, B. (2008) 'The Human Prejudice', in A. W. Moore (ed.), *Philosophy as a Humanistic Discipline*. Princeton, NJ: Princeton University Press.

Williams, J. (2015) ‘Public Reason and Prenatal Moral Status’, *The Journal of Ethics*, 19, pp. 23—52.

Williams, J. (2017) ‘War and Global Public Reason’, *Utilitas*, 29: 398—422.

Wolff, J. (2010) ‘Cognitive Disability in a Society of Equals’, in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Wong, S. I. (2010) ‘Duties of Justice to Citizens with Cognitive Disabilities’, in E. F. Kittay and L. Carlson (eds.), *Cognitive Disability and its Challenge to Moral Philosophy*. Oxford: Wiley-Blackwell.

Zimmerman, M. J. (1996) *The Concept of Moral Obligation*. Cambridge: Cambridge University Press.