

# University of Alberta

Problems of Extension in Justice as Fairness

by

David Pitcher

A thesis submitted to the Faculty of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Philosophy

©David Pitcher  
Fall 2011  
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

## **Abstract**

In John Rawls's liberal theory called "Justice as Fairness," citizens are conceived as reasonable and rational, and this conception of citizenship forms the basis for constructing principles of justice. Rawls notes that it is unclear how his theory of justice is to apply to entities other than citizens, and calls these cases "problems of extension." In this thesis, I discuss the way in which problems of extension arise in Rawls's work, and argue that there are underlying Kantian assumptions that lead Rawls to regard the conception of personhood as both necessary and sufficient conditions for being owed duties of justice. I argue that, on a constructivist interpretation of Rawls's theory, these assumptions are superfluous and threaten the political aspect of Justice as Fairness. I explore a revised version of Justice as Fairness wherein duties to non-citizens are acknowledged, and conclude that the problems of extension can be solved.

## **Acknowledgements**

I would like to thank my supervisor, Howard Nye, for his engaging, thoughtful and constructive comments and for the regular feedback on this thesis, and Jennifer Welchman, for encouraging me to consider an academic career in philosophy. I would also like to thank my father for helping me find my motivation, my mother for always taking an interest in my interests and proofreading my essays, and my brother who has been an unflagging source of inspiration.

## Table of Contents

Introduction.....	1
A Summary of Justice as Fairness .....	3
Problems of Extending Justice to Non-Citizens .....	16
Coercive Enforceability and the Domains of Justice and Morality .....	39
The Role of Personhood in Justice as Fairness .....	46
The Role of Constructivism in Justice as Fairness .....	53
Kant's Moral Theory and Respect for Autonomy .....	66
Coercively Enforceable Duties to Political Patients .....	81
Conclusion .....	96
Bibliography .....	98

## **Introduction**

In political philosophy, there are theories of justice that defend the idea that citizens in a democratic state should have equal and inviolable rights, as well as the means to exercise them. These theories – called liberal theories of justice – emphasize that beliefs and values generally differ from person to person, and that many of these values may be incommensurable since they often stem from many different religious or moral doctrines. However, if citizens can at least agree on a basic scheme of rights, then they should be able to cooperate despite the variety of their beliefs and values. Liberal philosophy invokes the concepts of right, equality, and freedom in order to find harmony in dissonance, without necessarily hushing the dissonance. Indeed, this ‘pluralism’ is itself a value in liberalism, and not merely because suppressing pluralism is a slippery slope to tyranny, but mainly because liberals see discord as a natural and inevitable result of the exercise of reason, without which citizens could hardly be said to possess autonomy. The values of autonomy, reason and reasonableness are central to liberalism, and are invoked to explain why we should bother getting along with people unlike ourselves.

In John Rawls’s liberal theory called “Justice as Fairness,” these values, and the rights they confer, are modelled on an ideal citizen, who cooperates fairly while exercising her rights in order to express and realize her reasonable beliefs and values. Since the ideal citizen is minimally defined by these reasonable moral and political attachments, the liberal state appears highly inclusive in that it does

not discriminate on grounds of race, gender, class, religion, etc. Those who are “excluded” are those who act unreasonably, say, by violating the agreed-upon laws or consistently infringing on the rights of others, but not because of arbitrary features such as skin color or genealogy. So, in this brief sketch of Rawls’s liberalism, our thoughts about justice emerge from the idea of a reasonable citizen, for the reasonable citizen.

In a political theory that emphasizes the value of reasonable autonomy and the role of the citizen, one might wonder about the theory’s treatment of non-citizen humans and non-human animals. How do they fit into liberal theory? What relations do they have to citizens, and what duties do institutions have to them? Is the purview of liberalism’s concept of justice wide enough to ascribe rights to entities other than autonomous and reasonable citizens? If so, which entities deserve rights, and how do these rights differ from the rights of citizens in a liberal state? Rawls recognized the difficulty of these problems, and referred to them as *problems of extension*. These are problems which are outside the scope of his liberal theory, but warrant our attention after working out a theory of justice for citizens in a Western democratic society.<sup>1</sup> In this paper, I intend to explore these problems in Rawls’s work. I will analyze the role of personhood and citizenship in Justice as Fairness and discuss the relation between these ideas and Kant’s moral theory. I will argue that Rawls’s insistence on the supremacy of typical adult humans in conceptions of citizenship and social membership traces back to Kantian foundations that are unsuitable for a *political* liberalism, and that

---

<sup>1</sup> Rawls, *Political Liberalism*, 20.

a suitably revised Justice as Fairness has the resources to include certain non-citizen humans and non-human animals as right-bearers in a liberal democracy.

### **A Summary of Justice as Fairness**

In Rawls's writings, there are two projects to distinguish. One of these projects is called "Rightness as Fairness," and the other is called "Justice as Fairness." The first is mentioned in *A Theory of Justice* as a moral theory that provides a measure of right and wrong conduct. The second, Justice as Fairness, is the political theory elaborated throughout the course of Rawls's career. In later works, such as *Political Liberalism* and *Justice as Fairness: a Restatement*, Rawls clarifies the difference between these two projects. Rightness as Fairness is an example of a "comprehensive doctrine," which is a theory that applies to "all subjects and covers all values."<sup>2</sup> While this definition may seem exceedingly broad, the main idea is that certain theories, especially "religious, philosophical, or moral" theories, offer principles, standards or guidelines that prescribe conduct for interpersonal affairs, lifestyle choices, sexuality, personal expression, community membership and so on.<sup>3</sup> They are generally wide enough in scope to assist someone in praising or blaming others, or choosing actions under everyday or novel circumstances.

When Rawls invokes the idea of a comprehensive doctrine, he has in mind religious doctrines that may or may not offer metaphysical explanations of the

---

<sup>2</sup> Rawls, *Justice as Fairness: A Restatement*, 14.

<sup>3</sup> Rawls, *Justice as Fairness: A Restatement*, 14.

world, and at least provide a system of norms and expectations that measure right and wrong, pious and sinful. He also has in mind several positions in moral philosophy that prescribe how to act, feel or judge in the majority of situations faced by anyone with a conscience. Exemplary of such a comprehensive philosophical doctrine is *utilitarianism*, the ethics that reduces right and wrong to the good and bad experiences of sentient creatures, and that typically prescribes the maximization of the good experiences (pleasure, preference satisfaction, etc.) and the minimization of the bad experiences (pain, preference frustration, etc.). Utilitarianism's comprehensiveness is particularly broad, since it can, in theory, apply to any possible situation that involves sentient creatures.<sup>4</sup>

An alternative to a comprehensive doctrine is a *political conception of justice*, which has a significantly more limited scope, since it does not directly apply to everyday choices and judgements, or to citizens and creatures, but it applies to the "basic structure" of society.<sup>5</sup> The basic structure is the set of political, social and economic institutions that coordinate and regulate the social, political and economic activity of citizens. Legislation, courts of justice, law enforcement, health care, education, and foreign policy are all examples of institutions that belong to the basic structure, which are directly affected by the political conception of justice. Furthermore, in order to accept or affirm a

---

<sup>4</sup> Utilitarianism and other comprehensive doctrines do not necessarily "apply to all subjects" in the sense that they provide prescriptions and descriptions for aesthetics, physics, chemistry, psychology, and so on. They apply to all "subjects" in the sense that fellow citizens are also fellow subjects in a state. A doctrine that, say, only applied to one's family members and pets would be a *partial* doctrine in this sense. The main idea behind a comprehensive doctrine *as comprehensive*, however, is that the values endorsed by the doctrine are ranked and ordered in a way that precludes its endorsers from agreeing with other such doctrines.

<sup>5</sup> Rawls, *Political Liberalism*, 11.



comprehensive doctrine, a given individual generally needs to possess specific beliefs and values. A political conception of justice, however, does not specify any unique grounds for its adoption and justification, since it must be able to be affirmed from a plurality of comprehensive doctrines. These doctrines “overlap” to form a collective affirmation. For example, a political conception of justice might include the idea that its citizens’ rights are inviolable and cannot be traded off for social benefits. A Kantian might agree to this idea by linking it up with the values of humanity and autonomy, a religious person might connect the idea up with certain divine commandments, and a utilitarian might see the practical benefits to be gained by regulating society with such a principle. Each of these people affirms the same idea, but from profoundly different bases. Lastly, while a comprehensive doctrine can invoke jargon and ideas that are not widely understood and recognized by the public, a political conception of justice must be expressed with and intelligible via common language and knowledge alone, similar to the way that public media avoids highly complex terms and esoteric notions in favour of familiar language and commonsense ideas. This collection of shared beliefs and values (that are reasonable) is called the “background culture” of Western democratic society,<sup>6</sup> and these are the beliefs and values that

---

<sup>6</sup> While Rawls focuses on constructing principles of justice from the background culture of a specifically *Western* democratic society, the main idea could presumably be applied to *any* democratic society. By specifying the society as “democratic,” Rawls hopes that the background culture will include ideas and values about individual autonomy and public decision making. By qualifying it as “Western,” Rawls wants to limit the conceptions of justice that are considered behind the veil of ignorance to those theories that have emerged from prevalent and long-standing traditions, such as utilitarianism and liberalism.

“overlap” to make a public consensus regarding the political conception of justice.<sup>7</sup>

In clarifying this distinction, Rawls emphasizes that Justice as Fairness should be understood as a political conception of justice, and not a comprehensive doctrine.<sup>8</sup> He sees the need to distinguish between these two kinds of doctrines because he assumes that the possible arrangements of regulation and cooperation for modern societies are limited by our present historical and social conditions. Following David Hume, Rawls calls these conditions “the circumstances of justice,” which are the conditions that have to obtain in order for a theory of justice to apply to a particular place and time.<sup>9</sup> These circumstances are represented by a list of economic, historical, social and psychological conditions that more or less describe the situation in the affluent sectors of our globe.<sup>10</sup>

---

<sup>7</sup> The beliefs and values that do not overlap can be reasonable or unreasonable, and they are presumed to belong to a comprehensive doctrine. If the “reasonable” beliefs and values are entirely absent from a given comprehensive doctrine, then it is not likely that any citizen who affirms that doctrine will also affirm the political conception of justice.

<sup>8</sup> Justice as Fairness could be affirmed from its comprehensive version, Rightness as Fairness, but, as a political conception of justice, it could never be justified by the comprehensive version alone. The comprehensive version of justice as fairness would become one of the many reasonable comprehensive doctrines that ultimately intersect to form overlapping consensus.

<sup>9</sup> Hume argued that the need for justice only emerges in societies with a scarcity of resources and a population that is generally peaceful, in which citizens tend to their own interests first: “The rules of equity of justice depend entirely on the particular state and condition in which men are placed.” (Hume, “Of Justice,” 91) Rawls has a more elaborate conception of the circumstances of justice, but the main idea remains constant – that the concept of justice simply does not apply where certain social and psychological conditions are not met.

<sup>10</sup> One might want to draw a distinction between circumstances of justice *in general*, which are the raw psychological and social conditions required for the concept of justice to apply, and the circumstances of *political* justice, which are the conditions required in order to achieve overlapping consensus. Rawls does not seem to distinguish between these two senses of the circumstances of justice, and often writes as if justice *just is* political justice. It’s unclear, for example, if Rawls would consider a feudal state just if that feudal state existed in an era with low life expectancy, high risk of disease and starvation, scarce resources, etc. Of course, Rawls is not committed to declaring such non-political states *unjust*; the concept merely fails to apply in those cases. It does, however, seem intuitive that one can talk meaningfully about justice and injustice outside of the context of Western democratic societies with reasonable pluralism, and that one

Because the principles of justice deal with resource distribution, there must be economic conditions of ‘moderate scarcity’ including limited resources, competitive acquisition, large ranges of wealth, and a lack of absolute poverty. In terms of historical conditions, the “fact of oppression” must be recognized, viz. that comprehensive doctrines can only be mandated by a state with illegitimate instruments of power, coercion, deception, fear and punishment.<sup>11</sup> This fact leads to the social conditions which include a democratic public that recognizes the “fact of reasonable pluralism,” which is the idea that people have and always will disagree with respect to their comprehensive doctrines to a reasonable extent. Finally, this social condition is the result of a deep psychological condition, labelled by Rawls as the “burdens of judgement.”<sup>12</sup> The idea is that there exists a tangle of barriers that impedes the exercise of one’s reason, including complexity of evidence, personal eccentricities when evaluating evidence, conceptual vagueness and impenetrability, and so on. These burdens of judgement, as psychological conditions that characterize our species, will be ever-present as both an explanation and forecast of the plurality that cuts across our beliefs concerning morality and religion.

Together, these circumstances of justice place limits on which doctrines can regulate the basic structure. A comprehensive doctrine is unsuitable because

---

could draw the distinction by using Hume’s criteria as criteria for circumstances of justice in general and Rawls’s criteria for circumstances of political justice. After all, our considered convictions about just and unjust regulations emerge from a world that is prior to the realization of political society.

<sup>11</sup> Rawls, *Justice as Fairness: A Restatement*, 84.

<sup>12</sup> While Rawls does not think every past or present society is characterized by reasonable pluralism, but that this condition is absent only in societies with oppressive governments, *because* of those oppressive governments.

enforcing one requires oppression, whereas a public conception of justice is suitable because its enforcement is affirmable from the many (reasonable) viewpoints that arise from “burdened” judgement. These limitations on suitable doctrines do not provide support for Justice as Fairness itself. They lend support to the idea that Justice as Fairness is the *right sort* of conception required to determine the basic structure, though it may turn out to be the wrong particular conception.

The support for Justice as Fairness comes from a thought experiment that is supposed to model our shared beliefs and values, especially those beliefs and values that are formed in light of reflection and consideration of alternative beliefs and values, without the pressures of haste, duress or other intervening forces. These beliefs and values are called *considered convictions*, and are, on Rawls’s account, the basic building blocks of any political conception of justice.<sup>13</sup> They are strongly-held opinions about political justice that are formed in circumstances conducive to good judgement, such as “tyranny is unjust, exploitation is unjust, religious persecution is unjust.”<sup>14</sup> These convictions are embedded in the background culture of a Western democratic society, and form the basis of overlapping consensus. The thought experiment – called the “Original Position” – is designed to argue for certain principles of justice that arise from these convictions, for the sake of overlapping consensus. To simplify matters, these

---

<sup>13</sup> In *Political Liberalism*, Rawls refers to these as “facts about justice” which can be discovered through history and implemented as fixed points in our thinking in political philosophy. (*Political Liberalism* 125)

<sup>14</sup> Rawls, *Political Liberalism* 124.

considered convictions guide us in knowing what is fair and unfair, and the Original Position tries to model this understanding of fairness.

In the Original Position, representatives of citizens come together to decide on the principles of justice that will govern the basic structure of society. The citizens themselves are not directly involved, but rather are represented by spokespersons who are working to secure the basic interests of the represented citizen. The purpose of using representatives is to eliminate personal biases, vendettas, or prejudices that particular citizens may have, and to ensure that the representatives themselves are not affected by special features of the represented citizen, such as the comprehensive doctrine to which he or she subscribes. In other words, the representatives are to secure the interests of citizens *qua* citizens, as defined by a conception of personhood<sup>15</sup> drawn from the considered convictions of a democratic society. A person, on Rawls's account, is defined by the following necessary and sufficient conditions, called the two moral powers:

“(i) One such power is the capacity for a sense of justice: it is the capacity to understand, to apply, and to act from (and not merely in accordance with) the principles of political justice that specify the fair terms of social cooperation.

(ii) The other moral power is a capacity for a conception of the good: it is the capacity to have, to revise, and rationally to pursue a conception of the good.”<sup>16</sup>

---

<sup>15</sup> Rawls also refers to the conception of personhood as a conception of “moral personality.” (*A Theory of Justice*, 329).

<sup>16</sup> Rawls, 18-19.

Together, these conditions respectively allow us to call citizens *reasonable* and *rational*. A citizen is reasonable if that citizen is capable of conducting oneself in accordance with moral considerations, such as the principles of political justice. Rationality, on this conception, is the twofold power of determining what one wants from life, and figuring out how to achieve it. In forming a conception of the good, a person must have the ability to reflect on and reason with values, and thereby determine for oneself which ends are worthy of pursuit. In pursuing a conception of the good with the power of rationality, a person must also be capable of applying the basic principles of “economic” rationality: performing simple means-to-ends calculations, choosing dominant or best alternatives wherever possible, consistently ranking final ends, and so on.<sup>17</sup> With these two moral powers in mind, the representatives of citizens in the Original Position strive to choose principles of justice that allow citizens to *be what they are*, as persons, insofar as they are rational and reasonable. In other words, principles that prevent citizens from exercising these moral powers are to be rejected at all costs.

This account of the necessary and sufficient conditions of personhood enters into Rawls’s argument for the principles of justice in a crucial way, because the principles of justice are intended to make it possible for citizens to exercise their two moral powers over a complete life, within the constraints of the circumstances of justice. The principles permit a citizen to exercise her capacities as a person, without dictating which conception of the good she should pursue.

---

<sup>17</sup> Rawls, *Justice as Fairness: A Restatement*, 87.

The account of personhood serves another purpose as well. It defines what representatives can consider about the citizens they represent, as well as what they *cannot* consider. Specifically, we imagine that the representatives are drawing up the principles of justice under a “veil of ignorance.”<sup>18</sup> The veil strips the representatives of any knowledge that they might have of their citizens’ contingent physical, social, economic and psychological features, such as class, gender, age, sexual orientation, conception of the good, and so on.<sup>19</sup> While the citizens will actually have rich cultural and personal identities, representatives behind the veil of ignorance can only conceive of citizens as entities with the two moral powers who are seeking to exercise them.

The veil of ignorance allows representatives to abstractly conceive of the interests of citizens, without actually knowing the *details* of those interests. For example, it might turn out that the represented citizen is a wealthy upper-class businessman, or a lower-class single mother. These different people will presumably have different interests because of their respective situations. The businessman, being wealthy, could afford private health care, and so might not have an interest in public health care, while the mother might need public support in order to afford medical care that is necessary to her children’s well-being as well as her own. However, since the representatives do not know whether their citizen is one or the other, they will aim to secure more general interests such as public health care, so far as it is to the benefit of the least-advantaged. In

---

<sup>18</sup> Rawls, *Justice as Fairness: A Restatement*, 87.

<sup>19</sup> Rawls, *Justice as Fairness: A Restatement*, 85.

accordance with what is called “maximin” reasoning, the representatives ensure that, in the case wherein the veil of ignorance is lifted and the represented citizen is among the least-advantaged, that citizen still has the best possible entitlement to goods (which is also sufficient to fulfill the citizens’ general interests.) The idea is to maximize the minimum amount of goods secured for the represented citizen and achieve the best result in the worst case scenario. Because each representative is focused on the *general* interests of the represented citizen behind the veil of ignorance, they can reach an agreement on principles of justice that secure those interests.

Rawls argues that the representatives behind the veil of ignorance ultimately settle on these two principles of justice:

1. Each person has the same inalienable claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all; and
2. Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least-advantaged members of society (the difference principle).<sup>20</sup>

In order to ensure that citizens are always able to exercise their moral powers, the first principle of justice is considered “lexically prior” to the second. This condition means that the first principle must always be satisfied before any consideration can be given to the second principle. In other words, institutions

---

<sup>20</sup> Rawls, *Justice as Fairness: A Restatement*, 42.



cannot deny rights to anyone, even if it would provide a greater benefit to the least-advantaged. The difference principle targets those who are the “least-advantaged” in the sense that they have the smallest lot of what Rawls calls “primary goods.” Primary goods are goods that people need in order to exercise their moral powers in general, regardless of their specific conceptions of the good, and they include financial resources, health care, education, and the “social bases of self-respect.” The social bases of self-respect emerge from the fact that citizens can regard each other as equals, and no citizen is granted more rights than any other citizen (and that these facts are common knowledge).

While the principles of justice emerge from the Original Position in order to ensure that citizens will have the means of realizing their abstracted moral capacities, the moral capacities become more detailed as the veil of ignorance is “lifted.” The veil of ignorance is lifted in four stages.<sup>21</sup> In the first stage, the veil is fully operative, and the principles of justice are chosen as laid out in the Original Position. In the second stage, the first principle determines the democratic constitution by outlining rights and freedoms. The constitution is drafted in the context of the representative’s newly-unveiled knowledge of the society’s “general facts” such as the society’s global location, affluence level, political climate, and so on.<sup>22</sup> In the third stage, more detailed legislation is created in accordance with the second principle to determine how institutions will distribute goods. The representatives gain access to general facts about the

---

<sup>21</sup> Rawls, *Justice as Fairness: A Restatement*, 48

<sup>22</sup> Rawls, *A Theory of Justice*, 197.

society's population, including statistics about size, education levels, minority groups, and so on, but the parties do not yet have access to the details of any specific citizens. In the final stage, the veil is *fully* lifted so that rules and regulations can be specifically applied to cases. At this point, the principles of justice are in full effect, and a citizen's sense of justice is defined as an affirmation of the principles of justice from that citizen's comprehensive doctrine. Citizens choose and pursue conceptions of the good "reasonably" in that they make use of their allotted rights and goods, within the bounds of their sense of justice, in cooperation with the ordinances given by the basic structure. If this is the case for all citizens, then overlapping consensus has been achieved, and the society is considered "well-ordered."<sup>23</sup>

Citizens affirm the principles of justice from within the context of their various comprehensive doctrines by a process called *reflective equilibrium*.<sup>24</sup> In reflective equilibrium, citizens try to mentally re-enact the Original Position using their personal set of considered convictions. In the re-enactment, one considers the representatives in the original position who are choosing between alternative sets of principles of justice. One reminds oneself that the representatives are subject to reasonable constraint, and these constraints are provided by one's considered convictions.<sup>25</sup> For example, with time and consideration, free from

---

<sup>23</sup> Rawls, *Justice as Fairness*, 5

<sup>24</sup> At the outset, citizens do not agree to any specific principles of justice, but they agree at least insofar as they collectively desire mutual terms of agreement that can form the basis of overlapping consent. This idea, which is the idea of a well-ordered society, motivates the citizens to undergo the reflective process.

<sup>25</sup> Rawls states that "Among [our] convictions are those about the restrictions to impose on reasons for favoring principles of justice for the basic structure, and these convictions we model

duress, haste and other confounds, one might come to believe that any form of slavery is wrong, even if one were to volunteer for enslavement. In imagining the initial choice situation, one would include the constraint that ‘no principle will be considered just if it permits or mandates enslavement.’ So, one imagines that the representatives dismiss principles of justice that, for instance, grant institutions the right to annul certain citizens’ rights to autonomy even if those citizens would benefit (in terms of, say, increased welfare) from their enslavement. So, in reflective equilibrium, one’s considered convictions become formal constraints for the decision procedure in the re-enacted original position.

There are several levels of reflective equilibrium: narrow, wide and full. If a single individual’s considered convictions, upon due reflection, support a conception of justice, then the equilibrium is deemed “narrow.”<sup>26</sup> If one engages dialectically with the conception of justice, by considering arguments for and against it, and one still ultimately affirms the conception of justice, then the reflective equilibrium is thought to be “wide.”<sup>27</sup> When all of the citizens within the democratic society affirm the same conception of justice out of their considered convictions, then a state of “full” reflective equilibrium holds. Finally, when a single political conception of justice is affirmed out of a multitude

---

by the idea of the veil of ignorance in the original position.” (Rawls, *Justice as Fairness: a Restatement*, 30). The considered convictions provide the content of the reasonable constraints which are applied to the representatives in the Original Position, and the principles of justice are said to “match” the considered convictions when the representatives can arrive at the principles of justice from that set of convictions. (Rawls, *A Theory of Justice*, 25).

<sup>26</sup> Rawls, *Justice as Fairness: A Restatement*, 30.

<sup>27</sup> Rawls, *Justice as Fairness: A Restatement*, 31.

of comprehensive doctrines, overlapping consensus is achieved as the ideal psychological infrastructure of a well-ordered society.

Now that the main concepts of justice as fairness have been summarized, the importance of the conception of personhood in the set up of the Original Position should be clear. The conception of personhood defines the veil of ignorance and helps the representatives understand what is at stake. After the principles of justice are implemented, the conception of personhood determines who will become the beneficiaries of the principles of justice once those principles are embedded in a constitution, embodied in legislation, and applied to institutions that govern the distribution of goods such as health care and education. The conception itself is regarded as a considered conviction, held in the background culture of a Western democratic society.

### **Problems of Extending Justice to Non-Citizens**

At the beginning of *Political Liberalism*, Rawls remarks that his “fundamental question” of determining the principles of justice for a well-ordered society composed of citizens in overlapping consensus leaves out a number of other important questions, the answers to which were postponed until the completion of the main project in order to simplify the question at hand.<sup>28</sup> Rawls refers to the unresolved questions as “problems of extension” in the sense that it is unclear how the scope of justice as fairness can be extended to apply to cases

---

<sup>28</sup> Namely, the question of how to specify the fair terms of agreement for a well-ordered society in spite of a reasonable pluralism of comprehensive doctrines. Rawls thinks that this question is the fundamental question of political philosophy.

besides fellow citizens. For example, without altering the original position, it is unclear how his conception of justice is to apply to “temporary disabilities and also permanent disabilities or mental disorders [...] future generations [...] the law of peoples [...] and] animals and the rest of nature.”<sup>29</sup> In the course of his work, Rawls provides answers to the issues of future generations and the law of peoples (i.e. foreign policy).

In *A Theory of Justice*, Rawls suggests that a useful answer to the problem of future generations can come from extending the veil of ignorance to exclude knowledge of the represented citizen’s generation – “since no one knows to which generation he belongs, the question is viewed from the standpoint of each and a fair accommodation is expressed by the principle adopted.”<sup>30</sup> In *Political Liberalism*, Rawls considers future generations as a problem of “just savings.”<sup>31</sup> In the original position, representatives ignore citizens’ specific generations but know that a political society involves cooperation over time, and therefore cooperation between generations.<sup>32</sup> So, they agree on a principle for savings that

---

<sup>29</sup> John Rawls, *Political Liberalism*, 20.

<sup>30</sup> Rawls, *A Theory of Justice*, 288. While a represented citizen could turn out to be in a future generation, it could not be the case that the represented citizen turns out not to exist at all, despite the vagueness inherent in not knowing *now* exactly who and how many will exist *then*. One of the stipulations of the original position is that the society in question has a population, and given general sociological facts, has a statistically projected population as well. In the original position, we imagine that citizens who compose the given population, as well as those who are projected to populate the society, are paired up with representatives, who are then placed under a veil of ignorance. Thus, the representatives can assume that the citizen they represent will exist sooner or later, and focus on the general nature of the represented citizen as defined by the two moral powers. Of course, it is possible that an operative liberal society could accidentally save for future citizens who turn out not to exist, due to, say, an unforeseen havoeking plague two decades hence that decreases the society’s birth rate, but these considerations do not affect the representative’s assumption that their citizen will exist.

<sup>31</sup> Rawls, *Political Liberalism*, 244.

<sup>32</sup> The sense of “cooperation” here is peculiar, because it seems that the present generation is always fostering the future generation while distant future generations are unable to benefit the

applies to all generations, in order to ensure that the principles of justice can apply from one generation to the next.<sup>33</sup>

The extension of Justice as Fairness to other nations is embellished into a full work, *The Law of Peoples*, where a modified original position produces principles of international justice that secure the fundamental interests of “peoples” or nation which may or may not be governed by liberal principles. Instead of individual persons being represented, entire peoples are represented in the modified original position. The veil of ignorance shifts to control for “the size of the territory, or the population, or the relative strength of the people[...] their natural resources, or the level of their economic development, or other such information,”<sup>34</sup> and eight principles of international justice are agreed on.

With respect to disabilities, Rawls believes that justice as fairness can offer a reasonable answer to the problem of extending the principles of justice from citizens to fellows who previously fully possesses the two moral powers, but have temporarily lost them. Because the parties in the original position are aware of general facts of psychology and sociology, they are aware that temporary disabilities are always a possibility, and so in the third stage of applying the principles of justice, parties will agree to a “reasonable law” that allocates funding

---

present generation in any significant way, besides, say, the hope of a continued heritage or some other benefit outside the scope of Rawls’s definition of primary goods. It seems that because backwards causation is impossible, there is no reciprocity between generations, and therefore no cooperation between them. However, the sense in which society involves cooperation over time and between generations is the sense in which we *trust* that, with the savings provided by the present generation, the future generation will affirm and apply the same principles of justice, maintain our just institutions, and continue following the principle of just savings for their own subsequent generation.

<sup>33</sup> Rawls, *Political Liberalism*, 273-4.

<sup>34</sup> Rawls, *Law of Peoples*, 32.

to the recovery of citizens who temporarily lack the two moral powers (i.e. entitling them to “normal health care.”)<sup>35</sup>

It is more difficult, however, to offer a reasonable answer to the case of permanent disability. Severe physical deficits and severe cognitive deficits both pose problems to Rawls’s account, but in different ways. In the case of severe physical disabilities such as quadriplegia or paraplegia, the citizen retains the level of moral personality required for entitlement to primary goods. However, it is unclear that the *same sort* of primary goods held by citizens without these severe disabilities is sufficient for the needs of quadriplegics and paraplegics. These people may be entitled to additional financial resources, and they may have the same access to medical and educational institutions in the sense that those institutions are equally available; nevertheless, quadriplegics and paraplegics may be unable to make full use of these resources by themselves, and they may require a significant degree of care in order to manage, understand and cope with their disability, especially while transitioning into the new lifestyle caused by reduced mobility and reduced ease of access to public facilities. As a result, their allotted wealth may not be *worth* as much as the wealth of their fellow citizens.

The case of severe cognitive deficits is also difficult for the principles of justice to handle. In these cases, it is unlikely that the person will ever obtain the full status of “personhood” such that they bear rights. If a permanent disability is induced by trauma after the person had already possessed the two moral powers,

---

<sup>35</sup> *Political Liberalism*, 245

he or she may fall under the previous reasonable law that entitled him or her to normal health care. However, he or she may altogether lose the status of personhood on account of the permanent disability, and his or her entitlement to welfare will be at risk. Furthermore, in cases where the severe cognitive disability is, say, congenital, it is unclear in what sense they are owed health care, education and other goods, or if there is a political duty to designate a caregiver or benefactor to assist them.

Of course, many cases of cognitive disability will not prevent the citizen from falling short of the two moral powers. The two moral powers are the minimum requirements for citizenship, and those requirements represent a very large range of cognitive capabilities. Nevertheless, not only must the person have a minimum degree of reason and rationality, they must have those powers to the extent that they can (at least in theory) undergo the process of reflective equilibrium and affirm the political conception of justice from within their comprehensive doctrines. There will inevitably be cases where a certain degree of reason and rationality is present, but this degree is insufficient for the person to engage in reflective equilibrium, and therefore insufficient to warrant the full host of rights and privileges accompanying citizenship. Because the liberal rights and privileges were designed to protect and respect citizens' autonomy as manifest by the two moral powers, there is a threshold of rationality and reasonableness, above which one is a person with protectable autonomy and below which one is merely a non-person human. VanDeVeer vividly presents the size of and variety within this group of non-person humans as follows:



“Suppose that only one percent of humanity fail to be moral persons on Rawls's criteria. We have, then, about forty million humans who are not, strictly, owed just treatment. Perhaps, these include, for example, anencephalic infants, the seriously psychotic, Tay-Sachs children, the quite senile, the irreversibly comatose, and possibly a subset of Down's Syndrome persons.”<sup>36</sup>

While some of these cases do not apply, these are the kinds of cases I have in mind that pose a problem of extension without a clear reasonable response. The first and last two cases may fall outside the scope of my discussion for different reasons. Anencephalic infants will never meet the criteria for personhood, but it is doubtful that they will meet another more minimal criterion for inclusion, on account of their insentience and virtual inability to survive birth. The irreversibly comatose or those in a permanent vegetative state will similarly be unable to meet any minimal criteria.<sup>37</sup> VanDeVeer acknowledges that the majority of Down syndrome cases will not pose a problem, since they “are usually able to learn self-help skills, acceptable social behaviour, and routine manual skills that enable them to be of assistance in a family or institutional setting.”<sup>38</sup> These cases, as well as many cases involving disorders associated with learning disability and mental retardation such as Turner's syndrome, Klinefelter's syndrome, would not prevent the persons in question from possessing the two

---

<sup>36</sup> VanDeVeer, 369.

<sup>37</sup> I am assuming that these more minimal criteria are suitable for a political conception of justice. There are moral criteria that are suitable for comprehensive doctrines that would prescribe special treatment for these cases, such as respect for the sacredness of human life. However, the idea of human life as sacred is unsuitable for a political criterion because it invokes language that is not publically comprehensible, and makes sense only on the context of a comprehensive religious doctrine.

<sup>38</sup> Butcher, Mineka, and Hooley. *Abnormal Psychology*. Thirteenth Edition. 587.

moral powers. Presumably, VanDeVeer has in mind cases of “profound mental retardation” that require “custodial care all their lives”<sup>39</sup> as in certain cases of microcephaly or hydrocephaly.<sup>40</sup>

As for the other cases that apply, I have in mind those cases where the mental disorder or cognitive disability impairs either reasonableness or rationality to such an extent that the individual fails to meet the threshold for personhood. VanDeVeer’s “seriously psychotic” presumably refers to a range of mental disorders that interfere with one’s moral judgement or rational capacities. For example, a catatonic schizophrenic’s reason and rationality may be compromised by erratic behaviour, uncontrollable impulses, idiosyncrasies in reasoning, and/or suicidal tendencies.<sup>41</sup> Also, a psychopath or sociopath would presumably fail to meet the requirements for a sense of justice, even though they may meet the requirements for rationality. In these cases, their sub-person status may make it unclear as to whether they deserve criminal punishment or psychiatric treatment, because these individuals will not be owed any duties at all by falling short of reasonableness.<sup>42</sup> Tay-sachs children and Alzheimer’s patients at later stages of cognitive decline may completely fall short of the two moral powers, and require intense and expensive care.

---

<sup>39</sup> Butcher, Mineka, and Hooley. *Abnormal Psychology*. Thirteenth Edition. 583.

<sup>40</sup> Butcher, Mineka, and Hooley. *Abnormal Psychology*. Thirteenth Edition. 588.

<sup>41</sup> Butcher, Mineka, and Hooley. *Abnormal Psychology*. Thirteenth Edition. 497.

<sup>42</sup> Prisons may offer psychiatric support, but not all provide counselling, and those that offer counselling may not offer the continual psychiatric care required in many cases. Butcher, Mineka and Hooley note that “mental health services are typically not provided for the majority of inmates who require them” (653).

One might suggest that a reasonable response to severe disability can be given by an analogy with the way that justice as fairness handles the case of children, but there is an important difference between cases of severe cognitive disability and the case of children. Because parties in the original position know that “political society is a system of cooperation from one generation to the next”<sup>43</sup> (since it is part of the fundamental idea of a well-ordered society), they can reason that there will need to be institutional support for children insofar as they are “future citizens.”<sup>44</sup> While children are not granted rights until they possess the moral powers (estimated by a legislated age of adulthood), as future citizens they are granted “claims.”<sup>45</sup> These claims are advanced by the adult members of their families. However, if those with severe cognitive disability are

---

<sup>43</sup> *Justice as Fairness*, 167n.

<sup>44</sup> *Justice as Fairness*, 157. One might question how far the “future citizenship” criteria can apply along the timeline of childhood development. If children are entitled to special support because of their future citizenship, then one might wonder whether there are other entities which eventually become citizens that are also worthy of special institutional support, such as neonates, foetuses, embryos, zygotes, ootids, sperm and eggs. Rawls presumably draws the line at “children” because children are psychologically complex enough to begin their induction into political society, whereas the previously listed entities do not have this degree of psychological complexity. Specifically, Rawls believes that the process of moral development begins at a child’s “morality of authority” and ends with the adult’s sense of justice (*A Theory of Justice*, 462). Because there is a “sequence of moral development” beginning at childhood and ending at adulthood, institutions which are concerned solely with autonomous entities nevertheless take an interest in fostering the moral capacities of children, and they recognize that “the necessity to teach moral attitudes (however simple) to children is one of the conditions of human life” (*A Theory of Justice*, 462). So, strictly speaking, the relevant aspect of children is not merely that they *will become* future citizens, but that they are on the psychological and moral path to *becoming* future citizens. In the case of infants who are too young to be regarded as beginning moral development, it is unclear what duties, if *any*, can be owed to them from the perspective of the two moral powers. Some other criterion is required to account for duties owed to infants who fall under this category.

<sup>45</sup> *Justice as Fairness*, 166. Rawls uses this language to describe the politically relevant interests of children. He says that the “claims of [parent’s] children as future citizens are inalienable” regardless of the situation (166). This ascription of claims to children might seem peculiar because children cannot advance those claims until they are full-fledged citizens. The idea is that we can anticipate their reasonable interests with the conception of personhood, and thus institutions such as education and health care will have a *duty* to nurture and educate children so that they develop reason and rationality (and whatever other skills and knowledge required to be a functioning citizen in political society). The claims of children are advanced for them by institutions and families.

granted any claims, it is not in virtue of any future citizenship. The legal backing to their claims must be footed elsewhere.

Another difficult case for justice as fairness is the case of non-human animals and the case of nature or the environment.<sup>46</sup> Rawls is especially uncertain that these cases can be given a reasonable answer, since he thinks that the solutions to natural or environmental issues are generally found from within comprehensive moral doctrines that are beyond the scope of political justice.<sup>47</sup> On Rawls's account, these matters are "not a constitutional essential or a basic question of justice, as these questions have been specified" by the fundamental question of how to achieve overlapping consensus.<sup>48</sup> While ultimately justice as fairness would put these matters up to public discussion, leading to a vote on environmental legislation, Rawls suggests that it is helpful, so far as it provides a clear answer of how citizens might act after the principles of justice are in place, to invoke the Christian idea that "animals and nature are seen as subject to our use

---

<sup>46</sup> The issues of cognitive disability and non-human animals may be tied together in some way. Elliot remarks that "There is no obvious non-arbitrary way of excluding animals from being represented in the original position that does not equally exclude some humans." (Elliot 101). In moral philosophy, Peter Singer also suggests that it is arbitrary or "speciesist" to consider the interests of the cognitively disabled and *not* the interests of non-human animals. (Singer, "A Utilitarian Defense of Animal Liberalism," 78). If this line of reasoning is correct, then if there is a solution to the problem of extending Justice as Fairness to the case of the cognitively disabled, then there might be a parallel solution for extending Justice as Fairness to the case of non-human animals. While Singer is talking about moral philosophy, the reasoning should apply *mutatis mutandi* to political philosophy as well. If the interests of the cognitively disabled are somehow relevant behind the veil of ignorance, then entities with similar psychological properties should be similarly relevant behind the veil of ignorance.

<sup>47</sup> *Political Liberalism*, 246.

<sup>48</sup> *Political Liberalism*, 246. In other words, Rawls is not suggesting that environmental issues have nothing to do with political philosophy *at all*; rather, environmental issues have little to do with his targeted question of determining the basis of overlapping consensus in a reasonably pluralistic democratic society.

and wont.”<sup>49</sup> By this, Rawls does not mean that citizens can needlessly harm animals, exploit the environment and extirpate species at their whim. Rather, he’s referring to the recreational, scientific and aesthetic values of nature, which will motivate citizens to draft legislation that protects special environments and species for the sake of present and future generations. So, the “claims of animals and the rest of nature”<sup>50</sup> are relevant to citizens at least insofar as their treatment of animals and nature benefits them as citizens. Rawls also suspects that we have certain considered convictions about non-human animals such as “it is wrong to be cruel to animals and the destruction of a whole species can be a great evil.”<sup>51</sup> These convictions “clearly [impose] duties of compassion and humanity” in many cases of non-human animals.<sup>52</sup> While Rawls emphasizes that this suggestion is speculation on how citizens will act after the principles of justice are in place, he thinks that it provides a satisfactory level of assurance, and that asking for more than this level is a slippery slope to being unreasonable.<sup>53</sup>

Many have responded that this solution is unsatisfactory. Pritchard and Robison argue that the stipulations of the original position cannot, by themselves, support Rawls’s suggested duties of compassion and humanity.<sup>54</sup> At the very

---

<sup>49</sup> Rawls, *Political Liberalism*, 245.

<sup>50</sup> *Political Liberalism*, 246

<sup>51</sup> *A Theory of Justice*, 512.

<sup>52</sup> *A Theory of Justice*, 512.

<sup>53</sup> *Political Liberalism*, 247. Rawls thinks that the Christian view has “the virtue of clarity and yields some kind of answer” but he does not, on these grounds, assert that the view is the definitive way of dealing with the case of non-human animals (*Political Liberalism* 245). He suspects that if we go further than this view, we would presumably be arguing from a specific comprehensive moral doctrine, and thereby “run the risk of being unjust, politically speaking.” (*Political Liberalism* 247).

<sup>54</sup> Pritchard and Robison, 56.

least, there is no *guarantee* of such duties.<sup>55</sup> They argue that Rawls cannot modify the veil of ignorance such that “the participants might end up, for all they could know, being animals once the veil is removed” because this modification would violate the Rawls’s ideal of reciprocity (since animals cannot agree or fulfill contractual obligations).<sup>56</sup> On their understanding of Rawls, the notion of reciprocity is essential to the argument for Justice as Fairness, such that the ability to engage in reciprocal relations with compatriots is a necessary condition for social membership in addition to the two moral powers. They continue to argue that the only possible modification of the original position that does not adjoin *ad hoc* premises (such as slipping in a utilitarian constraint that rejects principles that frustrate the preferences of sentient creatures) is that the parties in the original position simply impose mutual restraints on the treatment of animals and acknowledge duties of beneficence to them. Pritchard and Robison conclude that because the parties in the original position are ultimately *self-interested* with respect to one another, it would be arbitrary to assume that the parties have “compassion and humanity [...] only to animals and not to their fellow humans as well.”<sup>57</sup> They conclude that the Rawlsian framework has no resources to accommodate the duties to animals that Rawls suggests in *Political Liberalism*.

---

<sup>55</sup> Pritchard and Robison, 58.

<sup>56</sup> Pritchard and Robison, 60. If Richardson is correct that the ideal of reciprocity is entirely superfluous to the original position’s argument, then perhaps there is some way to expand the veil of ignorance to exclude “speciesist” criteria.

<sup>57</sup> Pritchard and Robison, 60. In *A Theory of Justice*, the parties in the original position are initially conceived as citizens and not representatives of citizens, and these citizens are using the original position as a fair bargaining situation so that they can secure the best lot of goods possible. Thus, the citizens are conceived of as ultimately self-interested. However, after *A Theory of Justice*, the parties in the original position are conceived of as representatives of

Others argue that modifications that extend the scope of justice as fairness to the case of animals are complicated, but nevertheless possible, from within the stipulations of original position. VanDeVeer argues that the difference principle requires *some* conception of the interests of sentient creatures – creatures that can have hedonic mental states such as pleasure or pain – otherwise the difference principle could not determine the scope of the “least-advantaged” members of society or who was “better-off” or “worse-off.”<sup>58</sup> He questions,

“If knowledge of race or gender or social position is excluded from participants [in the original position] in order to insure their neutrality in choosing principles to govern interaction among beings with morally relevant interests, must not species identification be excluded as well?”<sup>59</sup>

He concludes that a modified original position could result in principles of justice that recognize the interests of sentient creatures, including non-human animals and the cognitively disabled.<sup>60</sup>

---

citizens, and not the citizens themselves. The representatives understand that the citizens are not entirely self-interested, and the representatives are only interested in securing the citizens’ basic interests as a person. So, Rawls modifies the original position to make sure that citizens are not characterized as self-interested persons attempting to secure the best lot of goods. These citizens can be characterized as belonging to moral communities, political organizations, and having attachments to families, children, animals, etc. Pritchard’s and Robison’s argument applies best to justice as fairness as presented in *A Theory of Justice*, but the theory has evolved in ways that might harbour other resources for including the claims of animals.

<sup>58</sup> VanDeVeer 373.

<sup>59</sup> VanDeVeer 373.

<sup>60</sup> It seems to me that this argument confuses the interests considered in the difference principle with interests of a more specific kind. Rawls emphasizes that the purpose of the conception of primary goods in the Rawlsian framework is to define the least-advantaged. (Rawls, *Justice as Fairness*, 59). These primary goods are the abstract interests of citizens insofar as they are citizens with the two moral powers. Thus, when Rawls uses the language of “better-off” or “worse-off” he is referring to an index or measure as defined by the conception of moral personality, and not to, say, the basic interests or needs of a sentient creature in general. So, Rawls could acknowledge that we exclude species membership from being considered behind the veil of ignorance, but the

Brent Singer defends a version of VanDeVeer's argument. He argues that the complications of including non-human animals into the principles of justice is "no argument against it being morally right."<sup>61</sup> He considers Pritchard and Robinson's thesis that animals are ultimately excluded from Justice as Fairness, and attributes it to the claim that "Justice is a matter of reciprocity among moral beings only. Nonmoral beings fall outside the sphere of justice."<sup>62</sup> Without rejecting the ideal of reciprocity, Singer suggests that "reasoners in the original position and nonrational beings [such as nonhuman animals] *share* certain primary goods" such as "interests in potable water, safe air to breathe, and so forth."<sup>63</sup> He argues that to the extent that citizens in the original position identify with a "nonrational self," with nonrational interests, they could indeed "be incarnated as various animals" after the veil of ignorance is lifted.<sup>64</sup> It is unclear from Singer's argument, however, how tactic this evades the issue of reciprocity; it seems to be a more plausible answer to reject the ideal as altogether superfluous.

Robert Elliot takes a different approach. He argues that the account of moral psychology in *A Theory of Justice* explains the development of a citizen's

---

acknowledgement would not change anything at all. Only those with the two moral powers (regardless of species) would be granted citizenship and entitlement to primary goods. However, the set of *terrestrial creatures* with the two moral powers and the set of *humans* with the two moral powers are identical.

<sup>61</sup> Brent Singer, 224.

<sup>62</sup> Brent Singer, 225.

<sup>63</sup> Brent Singer, 226. While Singer's argument does not take into account Rawls's use of representatives of citizens in the original position, a similar argument can presumably be advanced *mutatis mutandi* for the case of representatives considering the nature of the citizens they represent behind the veil of ignorance.

<sup>64</sup> Brent Singer, 227.



sense of justice in a way that allows it apply to non-human animals.<sup>65</sup> If the sense of justice ultimately considers certain privations on the part of non-human animals as injustices, then presumably the parties in the original position will include reasonable constraints against harming animals in the principles of justice. He clarifies that the two main barriers in the Rawlsian framework to considering these reasonable constraints behind the veil of ignorance is that “the discussion is in thoroughly human chauvinist terms” and that “not all the primary goods [such as self-esteem] are essential to, or even helpful in, the successful execution of many non-human life plans.”<sup>66</sup> In other words, the conception of personhood guarantees that the original position will proceed in anthropocentric or human-oriented terms, and the index of primary goods that follows from this conception of personhood will be elaborate enough to advance the complex interests derived from the moral capacities of reasonableness and rationality. If the sense of justice really does include a sense of justice for nonhuman animals, Elliot thinks that the parties in the original positions can make judgements “from an animal’s point of view” without excessive error or absurdity, and these judgements will be sufficient for ultimately including nonhuman animals in the principles of justice.<sup>67</sup>

---

<sup>65</sup> Elliot, 100.

<sup>66</sup> Elliot 102.

<sup>67</sup> Elliot 103. Elliot’s argument might prove as impotent as VanDeVeer’s argument for the inclusion of animal interests if Elliot’s argument merely tries to conclude that species membership needs to be excluded from being considered behind the veil of ignorance. Specifically, if representatives behind the veil of ignorance disregard species membership, they still must focus on the two moral powers because they are still necessary conditions for consideration behind the veil of ignorance. Even if species membership is irrelevant, the two moral powers as necessary conditions will still lead the representatives to consider those with the two moral powers, regardless of species (viz. only reasonable and rational adult humans). The only difference would be that the word “human” never appears in the course of the original position’s reasoning. In order to see any change in the principles of justice, four broad changes need to occur to the

In “Rawlsian Social-Contract Theory and the Severely Disabled,” Henry Richardson argues that a “fundamental obstacle to extending the Rawlsian framework to the case of the severely disabled” is Rawls’s inclusion of an ideal of “reciprocity” in specifying the fair terms of cooperation between citizens in a well-ordered society.<sup>68</sup> Rawls suggests that the principles of justice should exemplify the idea that “all who do their part as the recognized rules require are to benefit as specified by a public and agreed-upon standard,”<sup>69</sup> and that, in the end, “the difference principle [is] a form of reciprocity.”<sup>70</sup> Richardson argues that this ideal is *extraneous* to the original position, and insofar as it is nonessential to the argument for justice as fairness, its central role in specifying fair terms of agreement can be reassessed.<sup>71</sup> Just as Rawls expands the veil of ignorance to exclude knowledge of one’s generation in solving the problem of extension in the case of future generations, Richardson suggests that the veil of ignorance can exclude knowledge of whether or not a represented citizen has a disability – “the parties are ignorant of the level of disability of the persons they represent.”<sup>72</sup>

---

stipulations behind the veil of ignorance. First, the conception of personhood needs to be regarded as *sufficient but not necessary*. Second, another sufficient-but-not-necessary conception needs to be included as one of the original position’s stipulations, such as criteria for being owed duties but not necessarily owing duties. Thirdly, the veil of ignorance needs to be changed so that representatives have no knowledge of whether their represented entity meets the criteria for owing duties or merely the criteria for being owed duties. Lastly, the list of primary goods needs to be changed in order to accommodate the “higher-order” or abstract interests of those who can be owed duties, in order to identify members of this class as among the least-advantaged members of society.

<sup>68</sup> Richardson, “Rawlsian Social-Contract Theory and the Severely Disabled”, 424.

<sup>69</sup> Rawls, *Justice as Fairness*, 6.

<sup>70</sup> Rawls, *Justice as Fairness*, 60.

<sup>71</sup> Richardson, 427.

<sup>72</sup> Richardson, 433.

With this method, discussion of the claims of the severely disabled is not relegated to the legislative stage of applying the principles of justice.<sup>73</sup> Rather, from behind the veil of ignorance in the original position, the parties can include among their “general knowledge” of psychology and sociology that “persons of all levels of ability and disability” are included, and that disability is a “pervasive and unavoidable feature of human life” resulting in a “continuum of disability.”<sup>74</sup> As a result, the two principles of justice are adjusted to include those with any level of disability. For example, the difference principle entitles the disabled to dignity and self-respect. In Richardson’s view, including the disabled behind the veil of ignorance breaks down the dichotomy between the disabled and the non-disabled in a way that greatly encourages the self-respect of the disabled.<sup>75</sup> In sum, Richardson demonstrates that, without the ideal of reciprocity, the original position can be modified to include the severely disabled by including relevant knowledge of disability in the general psychological and sociological knowledge of the representatives behind the veil of ignorance.<sup>76</sup>

Richardson’s alterations to the original position imply that Justice as Fairness, so long as it includes the reciprocity condition, will be unable to accommodate cases of physical disability wherein one still has the two moral

---

<sup>73</sup> Richardson, 432.

<sup>74</sup> Richardson, 442.

<sup>75</sup> Richardson, 446.

<sup>76</sup> Throughout the course of his article, Richardson considers Martha Nussbaum’s own solution to the Rawlsian framework’s inability to handle the case of the severely disabled. Her solution rejects the original position as a useful device for deciding between conceptions of justice and opts for a more Aristotelian variant, whereas Richardson assumes “the usefulness of Rawlsian social-contract analysis” (Richardson 423). Similarly, in focusing on the idea that the scope of justice as fairness can be extended, I also assume the usefulness of the original position as a clear and constructive way of rendering commonsense ideas into practical solutions to political problems.

powers, but is unable to provide contributions of the calibre required by the reciprocity condition. For example, while some quadriplegics may be able to contribute to political society, academia or fields that do not require manual labour, other quadriplegics that have invested their time and energy into a field that involves significant manual labour may find themselves unable to solitarily support their basic interests.

While Richardson thinks that there are resources internal to the Rawlsian framework that can handle difficult cases of disability, other philosophers such as Amartya Sen and Martha Nussbaum have concluded that there is no recovery for Rawls on the issue of the severely cognitively disabled, and that this failure poses a serious challenge to his account of justice.<sup>77</sup> Instead, these philosophers opt for an approach that is more *inclusive* (in the sense that it applies to more people than those with the two moral powers, including those with disability) and *substantive* (in the sense that it assumes a more value-laden conception of what is good for people in general). These approaches are called *functional* or *basic capabilities* approaches, because they focus on ensuring that citizens have all of the capabilities required to harness the goods that have been distributed to them.

The capabilities that Sen thinks a state should promote include *at least* health care, education, nutrition, survival, “political liberty and basic civil rights,”<sup>78</sup> freedom of transaction,<sup>79</sup> freedom to achieve,<sup>80</sup> entitlement to food,<sup>81</sup>

---

<sup>77</sup> Nussbaum thinks that Rawls cannot “handle this problem without a major overhaul of his theory.” (Nussbaum, 334)

<sup>78</sup> Sen, 15.

<sup>79</sup> Sen, 116.

low fertility<sup>82</sup> and mortality rates,<sup>83</sup> and public discussion.<sup>84</sup> Martha Nussbaum's approach invokes an even more specific list of basic capabilities, which she calls a list of "Central Human Capabilities."<sup>85</sup> This list represents a "social minimum" or a threshold of functional capabilities which all citizens are owed.

Sen worries that Rawls's conception of personhood and its corresponding list of primary goods grants citizens freedoms that are too formal. Citizens are entitled to the means to realizing their ends, but not all citizens may have the capability to realize these ends with the given means. For example, someone with a severe physical disability such as quadriplegia may be entitled to a minimum degree of primary goods such as education and health care, because she still has the two moral powers, but she may be unable to use these resources to the extent she requires with those powers alone. Resources that are crucial for living may be available but unusable to her without, say, the help of a caregiver or some other more specific assistance to help such people bring their means closer to their ends. In the case of severe *cognitive* disability, the individual may not even meet the criteria for the two moral powers, and therefore may be entitled neither to primary goods or caregivers.

---

<sup>80</sup> Sen, 117.

<sup>81</sup> Sen, 209.

<sup>82</sup> Sen, 199.

<sup>83</sup> Sen, 198.

<sup>84</sup> Sen, 226.

<sup>85</sup> Nussbaum, "The Capabilities of People with Cognitive Disabilities," 334.

Nussbaum is concerned that Rawls's account does not allow the cognitively disabled to be considered *at all* from behind the veil of ignorance,<sup>86</sup> and that the general psychological and sociological knowledge of the parties in the Original Position proves insufficient for considering the severely disabled as having "equal dignity."<sup>87</sup> The knowledge that these parties would possess would be merely descriptive, in the sense that they would know that people are overcome by temporary disability, and that others have permanent cognitive disabilities. However, they could not evaluate these facts *morally*, as Nussbaum's list of Central Human Capabilities does. Because the parties are restricted to valuing *only* moral personality, the claims of the severely disabled are relegated to being assessed only at the legislative stage of the application of the principles of justice, at which point there is still no guarantee as to how much consideration will be given to their interests.

To clarify, the parties in the original position cannot consider the claims of the severely cognitively disabled because the severely cognitively disabled simply fall short of having the two moral powers. The only place in Justice as Fairness where the interests of the severely disabled can be considered is after the veil of ignorance is lifted, after the constitution is drafted in accordance with the two principles of justice, and after the institutions that compose the basic structure are erected. At that point, the citizens with the two moral powers have been distributed their primary goods, and are able to actively engage in public reason, and work together

---

<sup>86</sup> Nussbaum, 334.

<sup>87</sup> Nussbaum, 335.

to pass legislation on specific but non-constitutional issues such as city infrastructure, pollution, and the treatment of the severely cognitively disabled. Because the severely cognitively disabled are only considered far into the application of Justice as Fairness, Nussbaum criticizes Rawls for altogether dismissing the severely cognitively disabled and having no resources with which to accommodate their interests.

If Nussbaum is correct that the severely cognitively disabled should be considered early in our theorizing about justice and about valid claims to entitlements, then the two moral powers included in the conception of moral personality cannot represent *both* necessary and sufficient conditions for being included in considerations leading to the principles of justice. In *A Theory of Justice*, Rawls himself suggests that the conditions of moral personality may be merely sufficient conditions for consideration behind the veil of ignorance.<sup>88</sup> However, in *Political Liberalism*, Rawls is confident that limiting the discussion to answering the fundamental question of justice allows him to “take the two moral powers as the necessary and sufficient conditions for being counted a full and equal membership of society in questions of political justice.”<sup>89</sup> In order to extend justice as fairness to cases other than people with the two moral powers, this statement should be interpreted as stating that the moral powers represent necessary and sufficient conditions for *full* membership, but they are not necessary conditions for *any* level of membership (for example, for being entitled

---

<sup>88</sup> Rawls, *A Theory of Justice*, 505.

<sup>89</sup> Rawls, *Political Liberalism*, 302.

to care and self-respect as a result of being explicitly considered behind the veil of ignorance). Nevertheless, it is unclear what sorts of alternative forms of membership are available in the society depicted in Justice as Fairness, and critics remain sceptical about the extent to which Justice as Fairness can accommodate the interests of various cases involving non-persons.

Rawls anticipated that these sorts of responses would follow from his speculation, and he qualified his comments on the more difficult cases as *preliminary*, stating that more serious consideration should eventually be given to those cases after the political project is completed. The reason for beginning justice as fairness with a conception of moral personality and then working outwards to consider problems of extension is that while justice among people is not *exhaustive* of the concept of justice, it is at least *typical* of it, in the sense that the clearest concerns of justice will involve issues between people. Furthermore, justice as fairness is rooted in a traditional theory in political philosophy called “social contract theory.” In social contract theory, the relation between citizen and state is regulated by a contract (or set of rules and obligations) to which citizens collectively agree or should agree, and thereby willingly concede some of their freedoms for the sake of mutual cooperation. Justice as fairness can be understood as a version of this “contractarian” philosophy, since, by virtue of reflectively playing out the hypothetical initial situation, the citizens retain their freedom despite subjecting themselves to the government’s coercive power. Lastly, this method is taken in order “to achieve a clear and uncluttered view of



what, for us, is the fundamental question of political justice.”<sup>90</sup> Because justice as fairness was originally intended as a response to the problem of finding a basis of agreement in the midst of reasonable pluralism, expanding the necessary and sufficient conditions for membership in political society would needlessly complicate the original position, whereas setting aside difficult cases for future discussion simplifies the main task.

This method of proceeding from personhood means that Rawls is working on one part of justice, and that any answers found there can be helpful in working through the other parts of justice. In other words, his conception of justice is part of a more comprehensive concept called *justice*, which is part of an even more comprehensive concept called *morality*.<sup>91</sup> The values in justice as fairness are ultimately “moral values.”<sup>92</sup> While Rawls does not define “moral values” outside of the context of specific comprehensive doctrines, he has in mind the sorts of values that are thought to be worth promoting for their own sake, by anyone who is capable of promoting them. Moral values are generally thought to be *universal* in the sense that everyone should endorse them, and are *overriding* in the sense that everyone should uphold them, regardless of any inclinations to not uphold them. They are invoked to justify one’s actions, and to pass judgement on the actions of others. Such values generally govern interactions between people, but they can also help one determine which actions and ends are worth pursuing in

---

<sup>90</sup> Rawls, *Political Liberalism*, 20.

<sup>91</sup> This is why Rawls agrees that liberal values can never be fully “neutral,” because they always trace back to moral values.

<sup>92</sup> Rawls, *Justice as Fairness*, 40.

life, in relationships and in difficult and sensitive situations. These values might also prescribe duties to those who can carry out those duties, and they may help one determine when one should feel remorseful or virtuous.

The moral values in Justice as Fairness include the value of moral persons and their moral powers, of permissible conceptions of the good, of a well-ordered society, and so on.<sup>93</sup> These moral values are at the same time deemed political because they derive from ideas “shared by citizens” that “do not presuppose any particular [...] comprehensive doctrine.”<sup>94</sup> These are the values that Rawls considers latent in the background culture of a given society, and they are the moral values that various reasonable comprehensive moral doctrines have in common. They belong to a larger set of moral values which includes other political and non-political values not outlined in Rawls’s conception of justice. If a problem of extension has a reasonable answer, it will stem from the sorts of values that are both moral and political. Such solutions cannot be *merely* moral, in that their values are not publically shared, because then the problem is solved in the wrong way.

In order to better understand what it means to extend the scope of justice as fairness to include other moral values that are also political, the sense in which Rawls thinks that justice is a subset of the broader category called morality needs to be clarified. In the next section, I will address the questions: what is the relationship between justice and morality in general? What is their relationship as

---

<sup>93</sup> Rawls, *Justice as Fairness*, 141.

<sup>94</sup> Rawls, *Justice as Fairness*, 141.

it concerns their respective domains or range of application? What distinguishes their domains? Answering these questions will help to clarify what it means to extend the scope of justice as fairness further into the domains of the broader concepts of justice and morality.

### **Coercive Enforceability and the Domains of Justice and Morality**

In this section I will describe Rawls's ideas that the domain of justice is a subset of the domain of morality, and that the domain of justice is circumscribed by those moral values which are *coercively enforceable*, in sense that the state can legitimately force its citizens to uphold them (by say, instituting punishments for failing to uphold them). In Rawls's language, a "domain" is a "family of [...] values."<sup>95</sup> In other words, a domain is defined as a cluster of values that share a property. For example, in justice as fairness, the "domain of the political"<sup>96</sup> is a cluster of values marked by reasonableness. Domains can overlap with other domains. For example, the domain of aesthetic values might overlap with the domain of values in a theory of environmental ethics, since a theory of environmental ethics may regard the aesthetic value of a certain natural phenomenon as morally relevant and prescribe the promotion of those values via the protection and preservation of the phenomenon.

One should not confuse the notion of a domain with that of *scope*.

Domains are sets of values with common features; the scope of a theory of justice

---

<sup>95</sup> *Justice as Fairness*, 182.

<sup>96</sup> *Justice as Fairness*, 182.

is the range of entities that are ascribed rights and privileges by its principles. For example, the first principle of justice mentions “each person,” and the second principle mentions the “least-advantaged members of society,” who are also persons. Therefore, the scope of the principles of justice is co-extensive with the set of entities with the property of personhood (within the well-ordered society).<sup>97</sup> The scope of a theory of justice is *limited by* the domain of values which it admits. If a theory of justice proceeded from the domain of “hedonic” values, such as “pleasure,” “preference-satisfaction” or “the absence of pain and preference-frustration,” then its scope would include all of those entities which are capable of hedonic mental states.

The view I wish to discuss is a view concerning *which* moral values are within the domain of justice. This view states that the values of justice are those that are reasonably “coercively enforceable.” In “The Minimal State,” Robert Nozick takes up and describes this view:

“Moral philosophy sets the background for, and boundaries of, political philosophy. What persons may or may not do to one another limits what they may do through the apparatus of the state, or do to establish such an apparatus. The moral prohibitions it is permissible to enforce are the

---

<sup>97</sup> The scope of the principles of justice is also different from their “subject.” The subject is the set of entities governed by the theory’s principles, which are the means of securing the rights and privilege of those within the principles’ scope. The subject of the two principles of justice is the basic structure of society, and Rawls believes that the scope is limited to those with the two moral powers (including those in future generations, and children who will come to possess those moral powers).

source of whatever legitimacy the state's fundamental coercive power has.”<sup>98</sup>

In other words, there are certain moral values that demarcate the domain of justice, and within that domain, the state (or the basic structure of a well-ordered society) is justified in enforcing its laws with coercive means, such as law enforcement, punishment, detainment, and so on. If it were not for those moral values, there would be no legitimate foothold for justice to impose its power on citizens. On Rawls's account, these coercively enforceable values are the *reasonable* values as defined by the political conception of justice.

Moral philosophy may also reveal many other moral values, but not all of these values are suitable for coercive enforcement. Some of these values are *private* insofar as they lack public support, since they must form part of a political conception of justice and derive from shared values in the background culture, and those who possess these private values and wish to realize them are compelled to use other avenues than the system of justice. For example, if a public official were to try and realize his private values by using the resources that come along with his position and power, this action might be deemed a conflict of interest. However, if this same official, in his personal time, were to try and realize his values by means such as lobbying or persuading others to join his cause, he would be permissibly pursuing his conception of the good in a reasonable manner. These *merely private* moralities are the values of comprehensive doctrines that do not overlap with the values of the political

---

<sup>98</sup> Nozick, “The Minimal State,” 19.

conception of justice, and the use of coercive enforcement to realize them is unjust.

Merely private moral values (and the duties which follow from them) do not exhaust the moral values which are unenforceable by the state's coercive apparatus. There are other classes of moral values which prove unsuitable for coercive enforcement, regardless of how widely accepted or publically affirmed these values are. They are generally those moral values that imply moral duties which are:

- (1) merely private, as already explained,
- (2) trivial, as recognized by common sense,
- (3) supererogatory or permissible for a citizen to fail to perform,
- (4) offense-preventing but not harm-preventing,<sup>99</sup>
- (5) impractical or costly to enforce,<sup>100</sup>
- (6) vague or complicated to the point of obscurity.<sup>101</sup>

It is unreasonable for a state to enforce any moral duties with these qualities. In a later section, I will discuss why each of these qualities is unenforceable from the standpoint of political justice, in order to distinguish between political duties which are coercively enforceable and merely moral duties which are not

---

<sup>99</sup> For example, there might be a moral duty to avoid offending others with one's clothing (say, by the depiction of offensive content), but unless one's clothing somehow happens to harm those around the wearer, this duty could not be enforceable.

<sup>100</sup> The issue of "partial compliance" falls under this category. Partial compliance is the situation where not all citizens obey a given law, and the noncompliance is tenacious despite law enforcement's efforts to encourage full compliance. When the government can expect only partial compliance with a certain law, and this partial compliance endures over time, the law may be considered impracticable and require revision or repeal.

<sup>101</sup> The line between impracticality and vagueness is not hard and fast, since a duty might be clear and impractical, but a duty cannot be vague and practical. Citizens cannot properly agree to vague or unduly complicated duties. A duty might be clearly defined and fully tractable, but entirely absurd to enforce. For example, one can imagine a clear duty to avoid wasting water when washing one's face, showering, doing the dishes, etc., and the duty can even specify how much water is to be allocated to certain washing acts. In this case, however, it is better to rely on the volition of citizens or to create regulations for water companies' fees and supply rather than cracking down on individual water wasters.

coercively enforceable, and thereby determine whether or not there are duties to nonpersons which are coercively enforceable in a political society.

On this view, the coercive apparatus of the state is also the arbitrator for disagreements between citizens. Citizens are to comply with a court's decision or face punishment. Nozick thinks that people are motivated to use courts of justice rather than their own coercive devices because the state's monopoly of coercive power is such that no citizen can impose his or her private values on another citizen – "the state does not *allow* anyone else to enforce another system's judgement."<sup>102</sup> Citizens advance their private values within the bounds of their *rights*, which are "enforceable claims" recognized by the system of justice.<sup>103</sup>

Nozick summarizes this understanding of the state as follows:

"The state is an institution (1) that has the right to enforce rights, prohibit dangerous private enforcement of justice, pass upon such private procedures, and so forth, and (2) that effectively is the *sole wielder* within a geographical territory of the right in (1)"<sup>104</sup>

In a word, the state has the exclusive privilege of enforcing the rights of its citizens. The state may be minimal and enforce very few rights, as in Nozick's own "libertarian" view. This state focuses on enforcing the rights to acquiring and transferring private property, and restores private property when these rights are violated by fellow citizens. The same sort of state may have a more robust system of rights, as specified by Rawls's first principle of justice, and it might use

---

<sup>102</sup> Nozick, 23.

<sup>103</sup> Nozick, "The Importance of Liberty and Self-Ownership," 676

<sup>104</sup> Nozick, "The Minimal State," 34.

its coercive apparatus for the restitution *and* distribution of property, with something akin to the second principle of justice. Either way, the underlying form of the state is such that its institutions enforce these rights, however simple or complex they may be, as outlined by a conception of justice.

In Rawls's discussion of "The Domain of the Political," he advances this view when he describes the political power of the state as "always coercive power backed by the state's machinery for enforcing its laws."<sup>105</sup> He emphasizes that this power is ultimately power granted by citizens, since the political conception of justice that regulates the state is justified by the overlapping consensus of citizens with a reasonable plurality of comprehensive doctrines.<sup>106</sup> Furthermore, citizens are viewed as "self-authenticating sources of valid claims" meaning that their claims are enforceable insofar as their claims derive from reasonable conceptions of the good, which is self-authenticating because it stems from each citizen's right to exercise the second moral power of rationality.<sup>107</sup> In this way, Rawls adopts the view that the granting of reasonable and enforceable claims is the business of the state.

To clarify, this view does not state that all there is to justice is the state's exclusive power to coercively enforce reasonable claims. In Nozick's view, the state is defined as having this power, and the view that I would like to argue for is

---

<sup>105</sup> Rawls, *Justice as Fairness*, 182.

<sup>106</sup> Rawls refers to this justification of the state's coercive power as the "the liberal principle of legitimacy." If this principle is not in full effect, then the relationship between citizen and state is merely one of obedience to authority, and the autonomy of each citizen is not reflected in the power of the state. *Political Liberalism*, 76.

<sup>107</sup> *Justice as Fairness*, 23.



that this aspect of the concept of justice is what circumscribes its domain within the larger concept of morality. This means that there is more to justice than coercive enforceability, viz. the properties that it shares with morality.<sup>108</sup> There are two claims here:

- (1) That justice and morality do not have distinct domains; political values are also moral values, and not all moral values are political values.
- (2) Within the concept of morality, the values that are both political and moral are those that are reasonably coercively enforceable. This latter notion circumscribes the domain of justice.

I have expressed this view in this manner to avoid presupposing an answer to the question, “what is the scope of these political values?” The idea is that claims are enforceable wherever they are reasonable, and the question “whose or which claims are to be recognized by the state” is left aside, to be answered separately. So, on Rawls’s view of the domain of justice, “extending the scope of justice as fairness” would mean examining the claims of a class of entities, linking those claims up with reasonable and coercively enforceable values, and then including representatives behind the veil of ignorance that act on behalf of those claims, adjusting the veil of ignorance accordingly to view certain characteristics of the bearers of those claims as arbitrary or irrelevant.

---

<sup>108</sup> Justice does not share all of its properties with morality, although its values should have the same properties as moral values. Thus, there are plenty of moral duties that are not coercively enforceable (such as those previously listed), but there are no political duties which are not enforceable.

## The Role of Personhood in Justice as Fairness

In the previous section, the question of “whose or which claims are to be recognized by the state?” was left aside. In this section, I will explore why Rawls thinks that the principles of justice only concern citizens, i.e. why Rawls thinks that the state can only legitimately coerce those with the two moral powers to uphold duties, and why duties can only be owed to those with the two moral powers.<sup>109</sup> I will begin by considering how Rawls thinks that a conception of personhood fits into his theory, viz. as a basic building block in the construction of a political conception of justice.

In *A Theory of Justice*, justice as fairness is understood as a contractarian model of justice. Rawls states that he intends to carry “to a higher level of abstraction the familiar theory of the social contract found, say, in Locke, Rousseau, and Kant.”<sup>110</sup> On the contractarian model, citizens agree (or ought to agree) to a “social contract” that specifies the terms of their cooperation. A state’s coercive power is legitimate when it is bound to such a contract (such as a constitution), and justice consists in the state’s strictly following this contract. Because of this agreement, citizens retain their autonomy even when their actions are restricted by the state or punished by its coercive power, because the state’s coercive power ultimately reduces to the power of its citizens.

It is clear how a conception of personhood fits into the traditional contractarian theory. Such a conception defines who is capable of entering into

---

<sup>109</sup> I will assume that these duties to citizens include duties to future citizens as well (as in the case of children and future generations, which do not pose strict problems of extension).

<sup>110</sup> Rawls, *A Theory of Justice*, 11.

contracts, who is able to fulfill contractual duties, and who is able to bear the autonomy that legitimates the state's coercive power. People are defined by their contract-making and contract-fulfilling capacities, and their mutual and autonomous agreement is the ultimate justification of a social contract.

The social contract may be actual or hypothetical. Actual social contracts or "social compacts" are used in the philosophy of Thomas Hobbes, wherein people emerge from an unsavoury "state of nature" or pre-contract world and are compelled by the drive of self-preservation to enter into contracts with their fellows. They eventually decide on a contract that will govern their affairs and mutually restrain their behaviour, and they collectively transfer to a ruler as many of their natural rights as necessary in order to uphold this contract. This ruler, or the "Leviathan," has the right to coercively enforce the contract and the responsibility to defend the nation.<sup>111</sup>

The trouble with actual social contracts is that citizens do not always give explicit consent to the social contract, and so consent has to be assumed on the part of citizens functioning in the contractarian society. This assumed consent on behalf of a citizen is called "tacit" consent, but it is not clear from the actions of a citizen who is making use of social goods (such as roads or markets) that they have tacitly signed a contract, or what binding force this consent would have even if it could be legitimately presumed. Furthermore, these contracts are not

---

<sup>111</sup> Hobbes, 608.

regulated by any notion of fairness, and so citizens could agree to social contracts that, for example, condemn themselves to enslavement.

In *A Theory of Justice*, Rawls views the contract defined by the two principles of justice as *hypothetical*, in the sense that, while there is no specific historical time and place where the contract was drafted, everyone can still *imagine* such a gathering of individuals and see the reasons they would have for being interested in such a contract. Instead of an actual “state of nature,” we have a hypothetical “original position” or “initial choice situation.” Any social contract we have here and now is justified only if its “principles would be accepted in a well-defined initial situation.”<sup>112</sup>

However, Rawls does not think that the use of this thought experiment amounts to full-fledged contractarianism or a “complete contract theory.”<sup>113</sup> He notes that while his theory uses a great deal of contractarian terminology, it is mainly to convey the idea that the principles of justice are justified if they would be chosen by rational persons.<sup>114</sup> A complete contract theory, however, would be unsuitable for a *political* conception of justice, since it would involve “the choice of more or less an entire ethical system” and would more likely be a comprehensive doctrine named “rightness as fairness.”<sup>115</sup> While the distinction between a political conception of justice and a comprehensive doctrine is not clarified in *A Theory of Justice*, it is ultimately what requires Rawls to drop the

---

<sup>112</sup> Rawls, *A Theory of Justice*, 16.

<sup>113</sup> Rawls, *A Theory of Justice*, 17.

<sup>114</sup> Rawls, *A Theory of Justice*, 16.

<sup>115</sup> Rawls, *A Theory of Justice*, 17.

idea of Justice as Fairness as a form of contractarianism, though it still uses the language of agreement.

In “Kantian Constructivism in Moral Theory,” Rawls shifts his understanding of the way that the original position justifies the principles of justice. Instead of appealing to the hypothetical agreement of rational persons in an initial choice situation in order to justify the principles of justice, Rawls thinks of the principles of justice as *constructed* out of shared moral premises, found at the intersection of various comprehensive doctrines in a Western democratic society. This common ground is composed of the considered convictions in the background culture of such a society. Rawls calls this new way of understanding the justification of Justice as Fairness “constructivism.”

Constructivism is a class of moral and political theories that invokes a metaphor of construction to describe, explain and justify a method of arriving at substantive conclusions. The sense in which these theories are “constructive” is twofold: first, they are constructive in the sense that they purport to work with basic materials to which everyone<sup>116</sup> has access, which are subsequently put together to form a coherent and stable structure. So, using a metaphor of construction, there is a procedure that specifies a building’s blueprints and architectural planning, and the considered convictions represent the raw materials out of which the building is made so that *we* – the builders – can get to work “here

---

<sup>116</sup> The scope of the term “everyone” is defined by the problem that a constructivist theory is attempting to answer. In Justice as Fairness, the problem is finding the basis of overlapping consensus in a reasonably pluralistic Western democratic society. Thus, “everyone” would refer to the members of a given reasonably pluralistic Western democratic society.

and now.”<sup>117</sup> In constructivism, we are the builders of morality but not its architects; we are not *radically creating* morality. Rather, we make use of what we have in our present state, and construct according to a fixed and predetermined methodology. Second, a constructivist theory is constructive in a pragmatic or practical sense – it is a “procedure that settles disputes.”<sup>118</sup> It sets aside issues of realism and relativism and proceeds from some suitable basis from which it constructs principles of morality or justice in the name of solving a problem at hand.<sup>119</sup> In Rawls’s case, the problem is finding the fair terms of agreement for a reasonably pluralistic Western democratic society, and the two principles of justice as constructed by the parties in the original position in order to solve this problem.

Rawls thinks that what makes his form of constructivism “Kantian” is that the most important material of construction used in Justice as Fairness is the conception of persons as reasonable and rational.<sup>120</sup> The principles of justice are “constructed” from this conception, as well as other non-specifically-Kantian considered convictions<sup>121</sup> such as “slavery is wrong.” When Rawls says that the

---

<sup>117</sup> Rawls, *Justice as Fairness: A Restatement*, 17. By the phrase “here and now,” Rawls is emphasizing (1) that we are trying to achieve reflective equilibrium in our present state, with all of our current intuitions and cultural assumptions and that (2) the decision procedure of constructivism is something that *we* actually *do*.

<sup>118</sup> Onora O’Neill, *Constructions of Reason*, (Cambridge: Cambridge University Press, 1989), 207.

<sup>119</sup> In Rawls’s words, “we try to avoid the problem of truth and the controversy between realism and subjectivism about the status of moral and political values. This form of constructivism neither asserts nor denies these doctrines.” Rawls, “Justice as Fairness: Political not Metaphysical,” 230.

<sup>120</sup> Rawls, “Kantian Constructivism in Moral Theory,” 516.

<sup>121</sup> In “Kantian Constructivism in Moral Theory,” Rawls also refers to these basic materials as “model-conceptions” and they are thought to include three main ideas: the well-ordered society, moral personality, and the original position (“Kantian Constructivism in Moral Theory,” 520). It sounds odd to consider the original position as a model-conception that is “latent in common

principles of justice are “constructed” from these considered convictions, he means that the reasonable constraints put on the parties in the original position are built in such a way as to model the important elements of our considered convictions. One of these reasonable constraints is the veil of ignorance. The conception of personhood, as a considered conviction, is transformed into a reasonable constraint, and the moral nature of persons is “represented” in these constraints.<sup>122</sup> In this way, the veil of ignorance is built to prevent the parties in the original position from considering anything besides the reasonable and rational nature of the persons they represent. They cannot, for example, consider the person’s gender, because it would violate the reasonable constraint provided by a conception of personhood which does not mention gender. Because these reasonable constraints are designed from the model of a reasonable and rational person, they ultimately provide constraints on the principles of justice that can be chosen, and in that way, the principles of justice also reflect the nature of persons as reasonable and rational. So, when Rawls says that the principles of justice are constructed out of our considered convictions, or that the principles “are constructed by justice as fairness as the content of the Reasonable,” he means that our considered convictions become reasonable constraints that are used in selection of the principles of justice from the original position.<sup>123</sup>

---

sense” (as all considered convictions are “latent” in the background culture of a Western democratic society). It might be better thought that the *values* of the original position are latent in common sense, such as the value of not discriminating on arbitrary grounds. The original position, as a full-fledged argument for the principles of justice, is far too complex to be considered a considered conviction.

<sup>122</sup> Rawls, “Kantian Constructivism in Moral Theory,” 529.

<sup>123</sup> Rawls, “Kantian Constructivism in Moral Theory,” 530, *Political Liberalism* 103.

Given how Rawls understands his theory, then, the conception of personhood has four roles in Justice as Fairness. The conception of personhood:

- (1) forms reasonable constraints on the selection of the principles of justice by restricting the parties in the original position to choice behind a veil of ignorance that narrows their vision to the reasonable and rational nature of the citizens they represent.
- (2) stems from the content of a basic considered conviction found in the background culture of a Western democratic society.
- (3) distinguishes Rawls's form of constructivism from other forms of constructivism, in that it makes it Kantian in nature.
- (4) provides the necessary and sufficient conditions for citizenship, or the necessary and sufficient conditions for right-bearing and duty-bearing.

It should be clear that the conception of moral personality, as a considered conviction and material of construction, provides an important premise for Justice as Fairness. The content of many concepts in Justice as Fairness can ultimately be traced back to the conception of personhood. The veil of ignorance has already been mentioned, and the list of primary goods is drafted in order to represent the "higher-order interests" of moral persons.<sup>124</sup> The principles of justice are also designed specifically to secure these interests. The conception of personhood shapes Justice as Fairness at every step in the construction of the

---

<sup>124</sup> Rawls, "Kantian Constructivism in Moral Theory," 526



principles of justice, and the nature of people as reasonable and rational is reflected in all the main concepts of Rawls's theory.

There are three main questions that I think are raised from the discussion in this section. First, how is constructivism – a doctrine that is ultimately about the status of moral truths – relevant to a political theory that specifies substantive political rights and duties? Second, how can a conception of personhood that emerges from Kant's comprehensive doctrine have such a prominent role in a political conception of justice? How essential is the conception of personhood to Justice as Fairness? I will explore each of these three questions in the following section.

### **The Role of Constructivism in Justice as Fairness**

I will begin this chapter by discussing the distinction between a normative theory and a metanormative theory in order to clarify the sense in which the two moral powers fit into the procedure of Justice as Fairness, as necessary and sufficient conditions that warrant coercively enforceable entitlements and duties. A theory is *normative* if it makes claims about what we have reason to do in some realm of life. For example, an ethical theory will make prescriptions that outline how we should and should not behave in a variety of situations. A normative theory may prescribe more than certain actions, since it may prescribe, say, how one should feel after witnessing an atrocity or how one should think about the legitimacy of slavery. Normative theories are not restricted to ethical theories, since they represent a broader class of theories. An aesthetic or epistemological

theory might tell us how to conduct art or identify knowledge, or what to deem 'beautiful' or 'objective.'

A *metanormative* theory, on the other hand, tells us about normative theories. Metanormative theories attempt to explain what it means to have reason to do such-and-such, what normative terms such as "ought" and "should" signify, and whether or not we should follow any normative theories to begin with. For example, a metaethical theory might assert that the moral prescriptions of certain moral theories are *true* because they relate to an independent metaphysical order of values. This sort of theory is called "moral realism" because it says that moral imperatives are expressions of metaphysically real values. Another theory, called "moral relativism," asserts that the moral imperatives are only "true" insofar as their truth is relative to a culture and its peculiar web of expectations, obligations and customs. Other theories claim that moral truth relates to one's subjective or psychological states. Expressivism asserts that moral prescriptions, judgements and imperatives are expressions of our attitudes, desires, preferences and other "non-cognitive" or affective mental states. Judgements of moral truth in expressivism are a matter of the moral judgement's aptitude for expressing these non-cognitive states. All of these metaethical theories describe *what it means* to pass moral judgement or to have true moral values, but they do not help us discern *which* moral theory is true (or if any are true at all.)

Thus, these kinds of theories are ways of understanding and explaining the phenomena of normative theories. A normative theory such as utilitarianism

invokes moral values such as “the Good” or “preference satisfaction” and derives substantive duties from these values.<sup>125</sup> Metanormative theories such as realism, relativism or expressivism can then be invoked to understand theories such as utilitarianism and explain the sense in which it uses terms such as “the Good,” how utilitarian reasons can be moral, and in what sense the prescriptions of utilitarianism are “true.” The metanormative theories are thought to be *independent* from the normative theories that they analyze. They are invoked to explain the meaning of normative terms in normative theories, but they do not immediately help us decide which normative theory is true, or whether any are true. Let’s call this independence of normative theories from metanormative theories “the metanormative gap,” since it refers to a gap between the two sorts of theories, such that, for example, substantive moral duties and judgements cannot be inferred from metaethical theories about the meaning of duties and judgements.

For example, in *Practical Ethics*, Singer holds that utilitarianism cannot be derived from metanormative considerations alone, and requires the substantive normative view that preference satisfaction is good and preference frustration is bad. Similar to Rawls’s constructivism, Singer argues that we can avoid the metaethical debate between realism and relativism. He argues that we should proceed from the fact that “ethical reasoning is possible,” and analyze the properties of ethical reasoning instead of analyzing the objective or subjective

---

<sup>125</sup> In the case of act-utilitarianism, there would be no list of duties, but rather a calculus is used to determine the moral action in any given situation.

state of ethical reasons.<sup>126</sup> So, he asks the metaethical question, “What is it to make a moral judgement, or to argue about an ethical issue, or to live according to ethical standards?”<sup>127</sup> In response to his question, he claims that our moral judgements are based on “ethical reasons” which are offered by people who are justifying their actions or lifestyles.<sup>128</sup> These ethical reasons must “be of a certain kind” in order to be ethical reasons at all, viz. they must be “in some sense universal.”<sup>129</sup> On Singer’s view, ethical reasons with a universal aspect must transcend an individual’s personal desires, preferences and goals, and consider those with similar desires, preferences and goals. Nevertheless, Singer does not think that these general claims about the nature of ethics “show that utilitarianism can be deduced from the universal aspect of ethics.”<sup>130</sup> Because of the metanormative gap, an independent normative theory must step in and argue that utility is the hallmark of an ethical reason, from which substantive duties and obligations can be derived.

Constructivism can act as a *merely* normative theory, without making any metanormative claims at all. As a normative theory, constructivism appeals to decision procedures such as the Original Position to generate substantive duties and obligations. As a metanormative theory, constructivism asserts that moral principles (such as the principles of justice) are “true” if they are the outcome of a constructive decision procedure (such as reflective equilibrium), or that to endorse

---

<sup>126</sup> Singer, *Practical Ethics*, 9.

<sup>127</sup> Singer, *Practical Ethics*, 9.

<sup>128</sup> Singer, *Practical Ethics*, 10

<sup>129</sup> Singer, *Practical Ethics*, 10-11.

<sup>130</sup> Singer, *Practical Ethics*, 14.

an moral principle *just is* to endorse the outcome of such a procedure. I shall describe Justice as Fairness as a normative constructivist theory, and then clarify the metanormative claims that it makes.

Justice as Fairness, as a normative theory, offers an account of the duties and obligations of citizens, institutions and states in a reasonably pluralistic political society, and provides reasons for these duties and obligations using the process of reflective equilibrium in conjunction with the idea of the original position. Justice as Fairness makes two claims that make it a normative constructivist view. First, it claims that the principles of justice are legitimate if and only if they are constructed from reasonable materials that are shared by people who affirm a variety of comprehensive doctrines and these materials are assembled according to a reasonable procedure. This is the sense of constructivism as *construction*. As a form of *Kantian* constructivism, the basic materials involve a conception of persons as reasonable and rational and a conception of a well-ordered society as mutual cooperation despite reasonable pluralism. Second, in accordance with the *pragmatic* sense of constructivism, it claims that no ideal solution to the problem of drafting the principles of justice for a well-ordered society is legitimate if that solution cannot be affirmed by or carried out by reasonable people. In other words, the principles of justice need to be affirmable from a variety of reasonable moral backgrounds, such that people holding different and conflicting comprehensive doctrines can all achieve wide reflective equilibrium regarding those principles of justice. The test of reflective

equilibrium acts as a measure to determine the extent to which the principles of justice are *pragmatically feasible* as a basis for overlapping consensus.

The first claim of Rawls's constructivism is made to establish that people in a well-ordered society *should* affirm the principles of justice, and the second claim is made to establish that people in a well-ordered society *will* affirm the principles of justice. To clarify the pragmatic sense of constructivism, Rawls's approach emphasizes that it is not enough to describe an ideal system that can coordinate political society; such systems need to take into account the extent to which people from various reasonable backgrounds will affirm the system. The just regime must not merely be *legitimate* in theory, but also *feasible* in theory.

The principles of justice are pragmatically feasible just in case citizens, exercising their rational and reasonable capacities in reflective equilibrium, can come to affirm the principles of justice despite the variety of their reasonable comprehensive doctrines. Recall that reflective equilibrium is at bottom the standard process of inquiry into moral and political philosophy, wherein one weighs the reasons for and against competing moral and political theories against their considered convictions without interfering factors such as haste and duress. At the end of this process, certain "objective" moral or political principles will emerge, but this objectivity does not necessarily refer to an independent moral order of values, to social or cultural artefacts, or moral intuitions that are revealed by reflection.<sup>131</sup> The sense in which these moral and political principles are

---

<sup>131</sup> Rawls, *PL*, 119.

objective is *only* the sense in which they have survived the process of reflective equilibrium. Furthermore, it might turn out that at the end of philosophical reflection, one discovers that, for example, utilitarianism is an objective moral theory. Thus, utilitarianism would be objective insofar as it survived *wide* reflective equilibrium. On Rawls's constructivist view, however, utilitarianism needs to survive *full* reflective equilibrium as well, wherein the state's population of citizens achieve wide reflective equilibrium. Even if utilitarianism is deemed objective from the perspective of wide reflective equilibrium, the main thrust behind Rawls's constructivism is that utilitarianism would be unsuitable as a candidate for full reflective equilibrium, since it is unlikely that all other reasonable comprehensive doctrines would come to affirm the substantive values of utilitarianism.<sup>132</sup>

This description of Justice as Fairness emphasizes only its merely normative characteristics. However, in "Kantian Constructivism in Moral Theory," and several sections of *Political Liberalism*, Rawls contrasts his merely normative constructivism with metanormative theories such as "rational intuitionism," which is the idea that moral judgements are true when they are entailed by our basic moral intuitions. He argues that Justice as Fairness, "as a freestanding view," requires that certain comprehensive moral doctrines *as well as metanormative theories* such as rational intuitionism are deemed unsuitable

---

<sup>132</sup> Because of the emphasis on pragmatic concerns, many forms of utilitarianism might come to affirm a liberal and political conception of justice. There are instrumental utility benefits to be gained, for example, by avoiding slavery, religious persecution, oppressive regimes and easily violable rights. At the very least, the utilitarian has *some* reason to affirm the principles of justice in reflective equilibrium, largely based on a shared concern for pragmatically feasible principles.

understandings of political conceptions of justice.<sup>133</sup> In addition, Rawls appears to be worried that citizens will not merely be reasonably pluralistic with respect to their various comprehensive doctrines, but that *they will be reasonably pluralistic with respect to their various metanormative doctrines* that make sense of the nature of the normative terms of the political conception of justice and their comprehensive doctrines. He states that:

“Many if not most citizens may want to give the political conception a metaphysical foundation as part of their own comprehensive doctrine; and this doctrine (I assume) includes a conception of the truth of moral judgements.”<sup>134</sup>

Rawls wants to ensure that these metanormative theories do not taint the legitimacy of the political conception of justice, and thus is motivated to replace “truth” with “reasonable” wherever moral truth is at issue. His goal is to say that Justice as Fairness, as a political conception of justice and merely normative theory, cannot have a metanormative foundation that prevents comprehensive doctrines from adopting their own metanormative foundations. Rawls argues that he is not rejecting metanormative interpretations of the political conception of justice; rather, many interpretations may be consistent with it, but the political conception of justice is as agnostic towards non-constructivist metanormative theories as it is agnostic with respect to the truth of any particular comprehensive doctrine.

---

<sup>133</sup> Rawls, *Political Liberalism*, 25.

<sup>134</sup> Rawls, *Political Liberalism*, 126.



Rawls adds a qualification to this claim, by saying that the political conception of justice cannot have a metanormative foundation *except* insofar as various citizens bring their own metanormative convictions in line with the political conception of justice in reflective equilibrium. However, the qualification seems trite or superfluous, since he has defined citizens as holding various reasonable comprehensive doctrines, but not as holding metanormative theories. In a society of squabbling philosophers, this qualification might be of paramount importance, but in the Western democratic society of Justice as Fairness, this qualification is largely irrelevant. What, then, does political constructivism's metanormative agnosticism have to do with coercive enforceability and the two moral powers?

The reason why Rawls makes these claims is that Justice as Fairness is only agnostic with respect to metanormative doctrines that specify the sense in which moral and political judgements are *true*. Rawls thinks, nevertheless, that a political conception must be tied to a particular understanding of *moral objectivity*, which specifies the conditions under which certain comprehensive doctrines are "correct" or "reasonable" as opposed to "true."<sup>135</sup> Thus, Rawls thinks that his normative constructivism requires a metanormative understanding of moral objectivity in order to be sure that the political conception of justice does not impose on reasonable comprehensive doctrines, while still being able to know

---

<sup>135</sup> Hill sums up the virtue of this approach as follows: "Rawls offers hope of a kind of moral objectivity missing in familiar emotivist and prescriptivist theories, such as those of Stevenson and Hare, while avoiding the strong metaphysical commitments of "moral realist" theories of the sort advocated by Plato, Moore, and others." (Hill, "Kantian Constructivism in Ethics," *Ethics* Vol. 99, No. 4 (The University of Chicago Press, Jul. 1989): p. 757.)

what makes full reflective equilibrium a viable candidate for political justification.<sup>136</sup> Rawls states that:

“Kantian constructivism holds that moral objectivity is to be understood in terms of a suitably constructed social point of view that all can accept.”<sup>137</sup>

In other words, the meaning of moral objectivity is specified by the “suitably constructed social point of view” which is the political conception of justice, viz. Justice as Fairness. In this way, Rawls thinks that political constructivism needs to make *both* normative and metanormative claims in order to distinguish itself as a fully “free-standing view,” and to distinguish between “true” and “correct” views.<sup>138</sup> In order to engage in reflective equilibrium to endorse the principles of justice, one must adopt a “social point of view” in the sense that one must believe that the results of the reflective process result in an objective or “correct” solution. Otherwise, there is the risk that one might merely adopt the principles of justice as

---

<sup>136</sup> The requirement that *some* kind of objectivity is necessary for a normative theory is still consistent with metanormative agnosticism, since the requirement is minimal and can be easily met by other metanormative positions such as expressivism and realism. While the requirement might rule out some metanormative theories that dismiss *any* sense of objectivity altogether, it does not require those undergoing reflective equilibrium to commit themselves to Rawls’s metanormative views.

<sup>137</sup> Rawls, “Kantian Constructivism in Moral Theory,” 519

<sup>138</sup> Thus, when Rawls says that “it is only by affirming a constructivist conception – one which is political and not metaphysical – that citizens generally can expect to find principles that all can accept” he means that the political conception of justice can only act as the basis of overlapping consensus if it understands that moral judgements are objective and correct when they are constructed by a reasonable procedure, and their objectivity *consists in* this the execution of this procedure. (*Political Liberalism* 97) While “we do not say that the procedure of construction makes, or produces, the order of moral values,” it still makes and produces the objectivity of its outcome.

a *modus vivendi* or a way of life that happens to be consistent with the political conception, but without having truly endorsed it.<sup>139</sup>

Thus, Rawls thinks that political constructivism requires some metanormative understanding of the objectivity of principles derived from reflective equilibrium, and he is thereby motivated to breach the “metanormative gap” that separates any substantive normative view from requiring a particular metanormative interpretation. The constructive process requires a specific metanormative interpretation insofar as its objectivity is understood as consisting in construction from a social point of view, but this interpretation is intended to be consistent with particular interpretations of the “truth” of the principles of justice.

These claims relate back to the substantive normative aspect of the theory, in that coercively enforceable principles need to be objective in this sense.

How does this tangent about the metanormative side of constructivism bear on the idea that the only coercively enforceable duties are duties to people, as defined by the two moral powers? Rawls’s metanormative claims provide a way of understanding the role of the considered convictions in the process of reflective equilibrium, and explains *why* the conception of personhood has the roles that it does (as the necessary and sufficient conditions for owing and being owed duties,

---

<sup>139</sup> Rawls, *Political Liberalism*, 126. A well-ordered society that finds support from various comprehensive doctrines based on the pragmatic feasibility of its principles might run the risk of its citizens affirming the principles of justice as a *modus vivendi*, since some citizens, such as utilitarian citizens, might affirm the principles of justice because those principles are pragmatic means to realizing their ends, so far as they acknowledge that not all other citizens will adopt their personal views. However, if such citizens have really affirmed the principles of justice *out of* their comprehensive doctrines, then they would not merely adopt the principles as a *modus vivendi* and try to impose their values wherever possible, but would see the pragmatic value of political liberalism as worth pursuing even from their personal comprehensive doctrine.

as a basic material of construction, etc.), i.e. why Rawls thinks that political constructivism must be *Kantian* constructivism. Specifically, Rawls states that the conception of personhood is needed in order to “work out the idea that the principles of justice issue from a suitable procedure of construction.”<sup>140</sup> In Rawls’s view, the objectivity of the principles of justice is made possible by the construction and reflection of the qualities of reasonableness and rationality included in the conception of personhood. Without the conception of personhood, there could be no clear content to the idea of the Reasonable – “we come to understanding this idea [of the reasonable] by understanding the two aspects of the reasonableness of persons and how these enter into the procedure of construction and why.”<sup>141</sup> The objectivity of the principles of justice is and must be constructed out of the conception of personhood. So, in setting up the Original Position, Rawls assumes that if the notion of moral personality were not implicit in the background culture of society, then there would be insufficient materials with which to build *any* principles of justice at all, at least principles that can achieve overlapping consensus.

Rawls’s constructivist explanation of the inclusion of the two moral powers explains why the two moral powers are sufficient conditions for being owed and owing duties which are coercively enforceable, and it also explains why the two moral powers are necessary for owing duties. It is commonsensical that only those who are capable of owing duties should be coercively enforced to

---

<sup>140</sup>Rawls, *Political Liberalism*, 94.

<sup>141</sup>Rawls, *Political Liberalism*, 94.

perform on those duties, and so people from a variety of backgrounds can understand and apply the idea that the two moral powers are necessary for owing duties. Using the two moral powers as sufficient conditions for being owed and owing duties also creates a workable conception of the Reasonable, since the moral powers can be worked into reasonable constraints such as the veil of ignorance, and the principles of justice can be modelled to secure the clearly-defined higher-order interests of persons. So, the two moral powers (and not some other criteria) play this role in construction in the same way that steel plays an important role in the construction of skyscrapers. Skyscrapers are not made out of wood because wooden materials would not support the weight of the desired structure. Similarly, the liberal conception of moral personality is required as a material to build the structure of overlapping consensus. Conceptions of personhood based on the soul, on Kantian autonomy, or any other metaphysical speculation would simply fail to hold up the structure; instead, a more liberal conception of autonomy that can be endorsed from a breadth of comprehensive doctrines is required to brace the structure.

Rawls's constructivist explanation, however, fails to explain why the two moral powers are necessary conditions for being owed duties. The idea that reasonableness and rationality are required for being entitled to rights or goods in a way that is coercively enforceable by the basic structure is not required in order to successfully set up and make sense of the construction procedure and the content of the Reasonable. If those conditions are merely sufficient for entitlement to rights and goods, then it's *at least* the case that citizens will receive

the *same* benefits from the principles of justice, and their fundamental interests will be secured. Presumably, Rawls includes this idea as an expression of an underlying assumption of the Kantian ideal of “respect for autonomy” that forms the basis of political society in Kant’s moral philosophy, wherein the two moral powers are necessary conditions for being owed duties because they set up the aspect of humanity and autonomy in virtue of which an agent is entitled to respect. Rawls’s assumption of this ideal is evidenced in the fact that the conception of personhood represents *both* necessary and sufficient conditions for being owed and owing duties, rather than having the two moral powers represent merely sufficient conditions for being owed duties, but still necessary conditions for owing duties. In the next section, I will explore this issue of Rawls’s inclusion of the Kantian ideal of “respect for autonomy,” and argue that the conditions of moral personality should be regarded as merely sufficient conditions for being owed coercively enforceable duties.

### **Kant’s Moral Theory and Respect for Autonomy**

In *Justice as Fairness*, everything hinges on the meaning of “Reasonable.” This notion limits the extent of acceptable pluralism by hiving off liberally problematic comprehensive doctrines. It also defines the sorts of constraints that can influence our reflective process on the course of endorsing principles of justice. It circumscribes the kinds of entities that justice regards as subject to duties and obligations, and ultimately determines which moral values are coercively enforceable. Many of the claims that Rawls makes are qualified by the

expression “insofar as it is reasonable.” In the previous section, I discussed the sense in which Rawls thinks that the content of the “Reasonable” is comprehensible entirely from within the constructive procedure but not from any “true” idea of the Reasonable, and that the two moral powers play a critical role in cashing out the meaning of the Reasonable, i.e. that the principles of justice are modelled after our image.

In this chapter, I will examine Rawls’s idea that reasonableness and rationality are necessary conditions for being owed duties. I will begin by clarifying an important distinction between moral agency and moral patienthood. In moral theory, different kinds of entities can have different kinds of moral status, which determine the extent to which they owe or are owed duties. Moral status can refer to either *moral agency* or *moral patienthood*.<sup>142</sup>

Moral agents are entities that are bearers of duties, obligations and responsibilities. Typically, moral agents are understood as psychologically complex, with the ability to reason, form intentions, to make plans, and so on. Moral patients are the entities to which duties, responsibilities and obligations are directly owed; i.e. those worthy of moral consideration by moral agents. Some entities are both moral patients and moral agents, such as neighbours in a community. Other entities might be *only* moral patients, since, e.g., no one would

---

<sup>142</sup> Light, Andrew and Holmes Rolston, “Introduction: Ethics and Environmental Ethics,” *Environmental Ethics: an anthology*, (Malden: Blackwell Publishers, 2003), 5.

expect babies to fulfill duties of benevolence, but we might have a duty not to act unjustifiably cruelly toward them.<sup>143</sup>

A similar distinction could be made in the context of political liberalism. If an entity is capable of owing duties, and the state can legitimately coerce the entity to carry out those duties, then that entity is a *political agent*. If an entity is owed duties from the state or political agents, and these duties are coercively enforceable, then that entity is a *political patient*. Given this distinction, another way of understanding the issue at hand is by asking if the two moral powers are necessary and sufficient conditions for political agency (i.e. for owing duties of justice), but only sufficient conditions for political patienthood (i.e. for being owed duties of justice). If so, then some of the problems of extension may be solved by extending justice to account for cases of political patienthood where there is no political agency. If this is correct, then the reasonable constraints that characterize the original position could be shifted to accommodate for the change.

Representatives behind the veil of ignorance, for example, would not know if the

---

<sup>143</sup> It is difficult to come up with an example of an entity that is only a moral agent. From the perspective of a child, the parent would be only a moral agent, and not a moral patient *to the child*, since a child cannot owe duties. Nevertheless, in the broader moral community, the parent is also a moral patient *to others*. Strictly speaking, the entity would have to meet the criteria for moral agency while failing to meet the criteria for moral patienthood. Since many theories place both criteria on a continuum of psychological capacities, with more advanced creatures as moral agents and less advanced creatures as moral patients, many theories imply that moral agents necessarily meet the criteria for moral patienthood. Nevertheless, because such a continuum is merely a device for representing the relative complexity of various psychological capacities, and all of the moral agents with which we are acquainted are *a fortiori* also moral patients, it is still logically possible to conceive of an entity that is *only* a moral agent. For example, it might follow from utilitarianism that a highly complex robot that lacks sentience but can still make informed decisions is only a moral agent. Also, Sidgwick suggests that, in a consequentialists framework, God cannot possess moral patienthood, since it is absurd and impious to think that “the existence of God is made better by our glorifying Him.” Because a state of affairs wherein God’s welfare is improved is impossible, he cannot be a moral patient. Henry Sidgwick, *The Methods of Ethics: Seventh Edition*, (Indianapolis: Hackett Publishing Company 1981), 115. If God has obligations, then it would follow that God is *only* a moral agent.



represented entity is a political agent or a political patient, but would nevertheless know the higher-order interests of political agents (the primary goods as presently defined) and the higher-order interests of political patients (derived by abstracting from the conception of political patienthood). Political patients, then, might have a right to care and might be owed duties of beneficence and non-maleficence, and an entitlement to a minimum degree of welfare. In order to achieve this conclusion, it is first necessary to show that not all political patients are political agents (those with the two moral powers).

I will begin by explaining why Rawls includes this condition in Justice as Fairness, and argue that the idea that all political patients are political agents is part of a restricted “Kantian Interpretation” of Justice as Fairness.<sup>144</sup>

In *A Theory of Justice*, Rawls suggests that Justice as Fairness is a “procedural interpretation” of Kant’s doctrine of respect for autonomy, in the sense that Rawls claims to adopt the procedure used by Kant in his moral philosophy. Kant’s procedure involved identifying, clarifying and validating a moral law that is found within agents who possess reason and rationality. This “supreme principle of morality”<sup>145</sup> is supposed to implicitly guide our moral thinking, and Kant believed that, by clarifying this moral law, we can learn to consistently apply this principle to any situation. Kant states that:

---

<sup>144</sup> The reverse claim, that all political agents are political patients, is also unwarranted. Suppose we define political agency in Rawls’s manner (by possession of the moral capacities), and political patienthood in a utilitarian manner (by the possession of sentience). In that case, it is logically possible to conceive of a robot that has cognitive states and not hedonic states. This robot would be required to fulfill various duties while lacking any entitlement to the benefits of justice. Such an entity would be *merely* a political agent.

<sup>145</sup> Kant, *Critique of Practical Reason*, 47.

It would be easy to show how common human reason, with this compass in hand, knows very well how to distinguish in every case that comes up what is good and what is evil, what is in conformity with duty and what is not, if, without in the least teaching it anything new, we only, as did Socrates, make it attentive to its own principle.<sup>146</sup>

So, Kant claims to elucidate a moral principle which operates in the moral reasoning of any rational agent, without introducing any moral principles that we do not already tacitly use.

Kant's elucidation of the supreme moral principle begins with the notion of the *will*. For Kant, the will is the "capacity [of a finite rational being] to determine itself to acting in conformity with the representation of certain laws."<sup>147</sup> So, finite rational beings form "practical principles" or mental representations of rules "that contain a general determination of the will" (which effectively are sentences with the word "ought," "should" or "must" as the main verb). One can choose to "determine one's will" in accordance with a practical principle in the sense that one keeps the principle in mind and chooses to act on that principle or some other practical principle. For every action, there is some practical principle that describes it. For example, one might consider lying to a friend in order to save that friend the discomfort of knowing some unsettling truth. The corresponding practical principle is "I should lie to my friend in order to such and

---

<sup>146</sup> Kant, *Critique of Practical Reason*, 58.

<sup>147</sup> Kant, *Critique of Practical Reason*, 78.

such” and the will is what ultimately allows us to decide whether to act on that principle, or some other principle (such as “I should not lie to my friend”).

On Kant’s account, practical principles are called *maxims* when they are subjective, in the sense that they hold for a single will.<sup>148</sup> A maxim holds for a single will when it specifies the agent to which the principle applies. For example, the principle that “I should not lie to my friend” is expressed in the first-person, so it only applies to me. When practical principle does not specify the agent to which it applies, and instead applies to all agents generally, then the principle is called a *practical law*. We can turn the practical principle in the example into a law by eliminating the first-person reference: “one should not lie to one’s friends.”

There are two other sorts of practical principles (or practical imperatives) for Kant: *hypothetical* (non-moral) ones and *categorical* (moral) ones.<sup>149</sup> A hypothetical imperative is one that applies to an agent because that agent has some specific desire or goal in mind and wishes to pursue it. In the case of desires, a hypothetical imperative might say “if one is to fulfill the desire to satiate one’s hunger, then one should eat some food.” In the case of specific goals, a hypothetical imperative might say “if one is to aim to learn how to play music, one should find a music teacher.” Categorical imperatives, on the other hand, do not presuppose any such desire or goal on the part of an agent. These duties apply to agents regardless of their inclinations and wishes, and are

---

<sup>148</sup> Kant, *Critique of Practical Reason*, 153.

<sup>149</sup> Kant, *Critique of Practical Reason*, 67.

therefore unqualified and unconditional. While there are many hypothetical imperatives (since there are many possible desires and goals one might have), Kant believes that there is only one categorical imperative, although it can have several formulations depending on our perspective. There is only one categorical imperative because there are no qualifications (such as those present in hypothetical imperatives) to distinguish different categorical imperatives. When Kant elucidates the categorical imperative for the first time, he formulates it as follows: “So act that the maxim of your will could always hold at the same time as a principle in a giving of universal law.”<sup>150</sup>

This formulation is a result of the following constraints on morality: it must be unconditional, universal, devoid of empirical considerations or particular objects of desire that might change its universal validity, and effective only at the level of the form of the will’s maxims.<sup>151</sup> These constraints are Kant’s conclusions about the nature of morality and the form that imperatives have to take in order to count as “moral,” and he believes that the categorical imperative can generate substantive moral duties instead of merely telling us about the nature of morality. Like Rawls, Kant dismisses the metanormative gap, and he thinks that we can learn something about the duties that we owe to one another by discussing the nature of moral terms. The underlying idea is that by uncovering the metaethical properties of morality, we can see what kinds of constraints are to

---

<sup>150</sup> Kant, *Critique of Practical Reason*, 164. Note that while Kant has many formulations of the Categorical Imperative, there is supposed to be only *one* Categorical Imperative underlying the common rational understanding.

<sup>151</sup> Kant, *Critique of Practical Reason*, 45, 155, 160.

be placed on ethical reasoning. Within those constraints, the phrase “anything goes” applies, and we can apply a “guess-and-test” procedure to spell out our substantive duties (by entertaining maxims and checking them against the categorical imperative). Kant’s approach has been criticized as being *too* formal and thereby unable to generate any helpful moral content such as our substantive moral duties. In other words, Kant is trying to make normative claims about our substantive duties when he is actually making only metanormative claims about the nature of morality and the will. Neo-Kantians such as Korsgaard also tend to reject the idea of a metanormative gap, and argue that we can derive duties such as respect for autonomy and humanity by analyzing the form of a moral law.<sup>152</sup>

For the sake of explaining Kant’s view, I will assume that bridging the metanormative gap is possible in cases where metanormative considerations entail reasonable constraints on moral reflection. For Kant, “[it] is impossible to think of anything at all in the world, or indeed even beyond it, that could be considered good without limitation except a good will.”<sup>153</sup> A *good* will is a will governed by the categorical imperative and not governed by hypothetical imperatives.<sup>154</sup> A person with a good will, then, acts for the sake of duty, as opposed to acting solely for the sake of his or her personal desires and aims. In Kant’s language, we can say that a person is virtuous when he acts out of respect for the moral law, and not merely in conformity with it. For example, suppose that the categorical

---

<sup>152</sup> Korsgaard argues that by learning about the nature of normative reasons, we can discover that not only do we owe obligations to other autonomous agents *in general* (Korsgaard, *Sources of Normativity*, 136) but we can establish *specific* duties. For example, Korsgaard thinks that we can determine the conditions under which suicide is permissible. (Korsgaard 162).

<sup>153</sup> Kant, *Critique of Practical Reason*, 49.

<sup>154</sup> Kant, *Critique of Practical Reason*, 55.

imperative requires someone to exercise their talents instead of allowing those talents to wither away. A person who is acting merely in conformity with the moral law might be realizing her talents solely for the purpose of attaining personal wealth or the adoration of others. A person, on Kant's account, would be morally praiseworthy and acting out of respect for the moral law if she understands and is motivated by the understanding that (1) her actions are the most efficient means to attaining her personal ends, as in accordance with hypothetical imperatives, and (2) the maxim by which she realizes her talents is prescribed by the categorical imperative as a universal law.<sup>155</sup>

When a will is actually motivated by and acts by these moral constraints, and thereby gives the moral law to itself (since the will "determines itself" in its formations of practical principles), the will is said to have *autonomy*.<sup>156</sup> However, if the will operates by hypothetical imperatives, the will is said to have *heteronomy*.<sup>157</sup> The praiseworthiness or blameworthiness of an agent always relates back to the agent's state of mind, intentions, desires and maxims which

---

<sup>155</sup> Rawls thinks that (1) is similar to the moral capacity for rationality and (2) is similar to the capacity for reasonableness: "Here again we distinguish[...] between the rational and the reasonable, a distinction that parallels Kant's distinction between the hypothetical imperative and the categorical imperative. [...] This priority expresses the priority of right; and justice as fairness resembles Kant's view in having this feature." Rawls, *Justice as Fairness: A Restatement*, p. 82.

<sup>156</sup> Kant, *Critique of Practical Reason*, 83.

<sup>157</sup> Kant, *Critique of Practical Reason*, 83. Even if one does not exercise one's will autonomously, one is still said to have the *capacity* for autonomy, since it is always possible for the will to act autonomously or heteronomously in the absence of actual autonomous conduct. (Kant 78) An entity with this capacity is said to have "humanity" which is to be respected even when that individual does not act autonomously. With this train of thought, Kant famously formulates the categorical imperative in terms of humanity: "So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means." Interestingly, in the Kantian framework, even if a person *never once* exercises his or her autonomy, that person will still have a superior moral standing in virtue of his or her humanity when compared to the moral standing of a severely cognitively disabled person or a non-human animal.

caused the action in the first place. If that state of mind reflected autonomy, then the agent is praiseworthy. If the state of mind reflected heteronomy, then the agent is not praiseworthy, and could be blameworthy if he or she violated the autonomy of another agent. Violations of autonomy occur when an agent does not act out of respect for the moral law and prevents himself or another agent from exercising his or her autonomy. For example, if I were to bind an agent to a chair against his will, then the bound agent would be capable of forming maxims that passed the universal law test, but incapable of carrying them out, and therefore I have handicapped his autonomy.<sup>158</sup>

From all of these ideas, Kant introduces the notion of a “kingdom of ends” which is a “systematic union of various rational beings through common laws.”<sup>159</sup> A kingdom of ends is a society composed of potentially-autonomous rational beings who are unified in their collective affirmation of the categorical imperative. Each agent, in passing laws for himself or herself via the categorical imperative test, passes laws for every other agent, since the categorical imperative only allows one to convert a maxim into a law if that maxim is universalizable. Citizens of the kingdom of ends pursue their various inclinations and goals within the constraints provided by the categorical imperative. These constraints are

---

<sup>158</sup> There is a way of understanding Kant that renders such violations of autonomy strictly impossible. Because moral laws only apply to moral agents so far as they can carry them out in their circumstances and with their present capacities and resources, a moral agent strapped to a chair would be performing all of his or her applicable duties at any given time, having a kind of vacuous or trivial autonomy. If this reasoning is correct, then the “violation” in strapping an agent to a chair cannot be located in his or her duties and duty-performance, but rather in the fact that one violates the autonomy of another *through* the categorical imperative, by failing to act out of respect for the moral law. This sense is still enigmatic, and the issue of whether or not autonomy is violable remains.

<sup>159</sup> Kant, *Critique of Practical Reason*, 83.

spelled out by members of the kingdom of ends who consider the categorical imperative and they generate laws that are in accordance with it. Thus, members of the kingdom of ends who do not pass these laws do not need to consciously undergo the categorical imperative process for every law, but if a citizen reflects upon the law, it must “be able to arise from his will.”<sup>160</sup> Lastly, because every citizen has an “inner worth” in virtue of his or her humanity, each citizen is said to possess a “dignity” which the kingdom of ends aims to secure and protect.<sup>161</sup>

In the kingdom of ends, duties are only owed to citizens, and only citizens can owe duties. To this point, Kant states that “morality, and humanity insofar as it is capable of morality, is that which alone has dignity.”<sup>162</sup> Dignity, which is grounded in the capacity for autonomy, and only dignity, makes a citizen “fit to be a member of a possible kingdom of ends.”<sup>163</sup> The idea of inner worth and dignity simply does not apply to any entities that are incapable of acting out of respect for the moral law, because worth is defined as a possibly-autonomous relation of a will to the categorical imperative. So, non-agents such as non-human animals lack moral standing on Kant’s account, but there still may be duties to them because of their relation to agents. For example, we might have duties of compassion to non-human animals, because and only because an agent’s compassionate treatment of non-human animals might help to foster the habits of virtuous behaviour owed to other agents.<sup>164</sup> Strictly speaking, however, these

---

<sup>160</sup> Kant, *Critique of Practical Reason*, 84.

<sup>161</sup> Kant, *Critique of Practical Reason*, 84.

<sup>162</sup> Kant, *Critique of Practical Reason*, 84.

<sup>163</sup> Kant, *Critique of Practical Reason*, 85.

<sup>164</sup> Kant, *Critique of Practical Reason*, 564



duties are not *owed* to non-agents – they are owed to fellow agents or to oneself – and therefore non-human animals and other non-agents cannot have the moral standing of patienthood on Kant’s account. Because of Kant’s doctrine of respect for autonomy and dignity in humanity, the kingdom of ends includes the moral powers as both necessary and sufficient conditions of membership in the moral community, and all moral patients are also moral agents.

Rawls believes that Justice as Fairness has its roots in Kantianism.

Rawls’s interest in Kant can be seen in his early writing, such as in “The Sense of Justice” where Rawls argues that the sense of justice can be paralleled with the Kantian sense of duty. He states that “one may follow Kant in holding that a good will, or in the present case, a sense of justice is a necessary condition of the worthiness to be happy.”<sup>165</sup> Here, Rawls is using Kantian language to express the idea that the moral powers are necessary conditions for patienthood. These ideas are finally unified in *A Theory of Justice* and are called “the Kantian Interpretation” of Justice as Fairness. There, he claims that “[t]he principles of justice are also categorical imperatives in Kant’s sense,”<sup>166</sup> and that the “veil of ignorance represents Kant’s understanding of morality as free from empirical and heteronomous considerations.”<sup>167</sup> While these claims lead one to believe that Rawls is identifying Justice as Fairness with Kantianism, in “Kantian Constructivism” Rawls emphasizes that “the adjective ‘Kantian’ expresses

---

<sup>165</sup> Rawls, “The Sense of Justice”, p. 304.

<sup>166</sup> Rawls, *A Theory of Justice*, p. 253.

<sup>167</sup> Rawls, *A Theory of Justice*, p. 252.

analogy and not identity.”<sup>168</sup> The analogy is drawn between Kant’s and Rawls’s respective procedures.

Rawls interprets Kant’s procedure as the construction of the categorical imperative from the conception of autonomous will. Rawls thinks that this construction is equivalent in form to using a conception personhood as the basic materials for constructing the principles of justice. Recall that “constructing” principles out of a conception of personhood involves transforming the qualities of persons into “reasonable constraints” that limit our thinking in the process of reflective equilibrium. Rawls says that “In a Kantian [constructivist] view, the conception of the person, the procedure, and the first principles must be related in a certain manner,” viz. the conception of personhood forms reasonable constraints that ultimately determine the content of the principles of justice.<sup>169</sup>

The basis of the analogy rests in two claims of Justice as Fairness:

1. The conception of persons as reasonable and rational is included alongside the idea of a well-ordered society in the basic materials of construction.
2. The procedure of construction uses reasonable constraints that are modelled using the conception of personhood. The veil of ignorance limits representative’s knowledge of the represented citizen to their

---

<sup>168</sup> Rawls, “Kantian Constructivism in Moral Theory,” p. 517; Also, see *Justice as Fairness: Political Not Metaphysical*, p. 224 – “the adjective “Kantian” indicates analogy not identity, that is, resemblance in enough fundamental respects so that the adjective is appropriate.”

<sup>169</sup> Rawls, *A Theory of Justice*, 517

moral powers; the aim of the Original Position is to secure the interests of citizens *qua* citizens with the two moral powers.

In Rawls's political philosophy, he limits the Kantian analogy to these two claims, so that he does not import Kantianism as a comprehensive moral doctrine.<sup>170</sup> The idea is to steer clear of Kantian comprehensive values such as 'respect for autonomy' or the value of human dignity, since these values cannot reasonably be shared in common by a reasonably pluralistic society. However, I shall argue that Rawls's inclusion of the claim that the two moral powers are necessary conditions of political patienthood (of being owed coercively enforceable duties) is based in the Kantian ideal of "respect for autonomy" and is therefore politically unfit to stand among the basic materials of construction. For ease of reference, I will refer to this requirement as the "Kantian requirement."

As already mentioned, there is evidence in Rawls's early writings in "The Sense of Justice" and the section on the "Kantian Interpretation" in *A Theory of Justice* that Rawls wishes to ground the claim in Kantian premises. In *Political Liberalism* and *Justice as Fairness: A Restatement*, Rawls takes "the two moral powers as the necessary and sufficient conditions for being counted a full and equal member of questions of political justice,"<sup>171</sup> and considers this assumption an "underlying philosophical conception" or basic considered conviction.<sup>172</sup> As

---

<sup>170</sup> "The comprehensive liberalisms of Kant and Mill, while viewed as suitable for non-public life and as possible bases for affirming a constitutional regime, are no longer proposed as political conceptions of justice." Rawls *Justice as Fairness: A Restatement* 198 Rawls purposely imports Kantianism as a comprehensive doctrine in *A Theory of Justice*, especially §40.

<sup>171</sup> Rawls, *Political Liberalism*, 302

<sup>172</sup> Rawls, *Political Liberalism*, 304.

clarified in the previous section, Rawls has explained that the reason he adopts this assumption is because he thinks it does the best job of providing content to the Reasonable, making constraints and resulting in a solution that can become the basis of overlapping consensus. While it seems undeniable that the two moral powers are sufficient for political membership, it is unclear why their necessity is warranted from Rawls's pragmatic justification. If the two moral powers *as sufficient* do the conceptual work required of them, then there is no further reason to make the two moral powers necessary conditions in addition to sufficient conditions. The two moral powers, as merely sufficient conditions, are nevertheless capable of performing their practical role: nothing in the argument for the original position would change in a way that compromises the higher-order interests of citizens with the two moral powers.

The best explanation, then, of why Rawls includes the Kantian requirement among the basic stipulations of the original position, is that it is a remainder of his early hyper-Kantian approach to political theory, viz. the position that respect for autonomy *completely* exhausts the grounds and ultimate ends of morality. The requirement can be found in Kant's idea of the kingdom of ends, and Rawls explicitly links the requirement to Kant's moral philosophy in his earlier writings.<sup>173</sup> The tenacity of the requirement in his later writings does not

---

<sup>173</sup> In *A Theory of Justice*, Rawls states that "the principles of justice are not derived from the notion of respect for persons, from a recognition of their inherent worth and dignity. Since the original position (as I have defined it) does not include this idea, *not explicitly anyway*, the argument for justice as fairness may be thought unsound." (*A Theory of Justice*, 585-6, my emphasis) Rawls indicates here that respect for autonomy may implicitly underlie his theory of justice, but his casual mention of it here means that he presumes that this implicit assumption is innocuous.

reflect Rawls's shift towards constructivism, but threatens to admit Kantian ideals of respect for autonomy and human dignity into the fundamental premises of the original position. We can then eliminate this requirement, on the basis that it has no constructive value and that it threatens to violate the politicality of Justice as Fairness.

The elimination of the Kantian requirement allows the political theory to have a greater scope, such that it may be able to handle the problems of extension as specified in an earlier section. It allows us to separate political agency from political patienthood, and consider candidates for the criteria that will define the political conception of patienthood as alternate sufficient conditions for membership. In the next section, I will discuss the sense in which the concept of political patienthood admits coercively enforceable duties in a way that solves certain problems of extension, and then respond to a worry that our duties to patients may not be politically relevant.

### **Coercively Enforceable Duties to Political Patients**

In this section, I will examine the notion of political patienthood by considering which conceptions of patienthood are permissible in Justice as Fairness's liberal framework. A conception of patienthood is like a conception of personhood, in that the conception of personhood defines the conditions for *owing* duties, and the conception of patienthood defines the conditions for *being owed* duties. A political conception of patienthood must spell out the minimum degree of psychological complexity required to be owed duties of justice, as well as the

need to invoke psychological features as such criteria. In order to have coercively enforceable duties to patients as opposed to merely moral duties to patients, the political conception must be a *reasonable* conception. According to Rawls's theory as expressed in an earlier section, a reasonable conception belongs to the background culture of a Western democratic society,<sup>174</sup> is "congenial" to the deepest convictions and traditions of that society,<sup>175</sup> and is "latent in common sense."<sup>176</sup> In other words, the conception must be simple and clear enough for citizens to understand, and it must have some strong connection to the academic heritage of the society.<sup>177</sup> For example, one of the reasons why the conception of the two moral powers is used to specify political agency is that the conception of the two moral powers is a broadly intelligible notion that is rooted in the liberal philosophy of Kant, Rousseau and Mill.<sup>178</sup>

The discussion of which capacities are best suited to define moral patienthood is a complex and controversial discussion. It would be best to follow Rawls's constructive method and suggest a working definition of moral patienthood in order to show how Justice as Fairness could make use of this

---

<sup>174</sup> Rawls, "Kantian Constructivism in Moral Theory," 569.

<sup>175</sup> Rawls, "Kantian Constructivism in Moral Theory," 569, *Political Liberalism*, 300.

<sup>176</sup> Rawls, "Kantian Constructivism in Moral Theory," 520.

<sup>177</sup> It seems that Rawls assumes that his set-up of the liberal society precludes the possibility of societies with wild and counterintuitive academic traditions. With the circumstances of justice – most notably the burdens of judgement – Rawls wants to assume that absurd traditions would be weeded out in favour of societies with a diversity of philosophical traditions with a minimal degree of credibility. Even if these traditions are wildly false, Rawls hopes that the amount of reflection and philosophical humility required of acknowledging the burdens of judgement still counts these traditions as *reasonable*.

<sup>178</sup> The other reasons why this conception is used as a premise in the original position is that it has constructive value, since it helps to organize our thought about justice and is capable of producing practical and substantive duties at the end of philosophical reflection. Furthermore, the intelligibility and historical aspect of the liberal tradition are intended to provide an initial reason for taking it seriously, and seeing how it works into a full theory of justice.

notion to solve the problems of extension. There may be other more suitable definitions of moral patienthood, or other sufficient conditions which help us recognize those who are political patients. This discussion, however, is outside the scope of this thesis. It suffices to show that there is *some* way of cashing out political patienthood, and that justice as fairness has access to conceptual resources that do not limit our coercive enforceable duties to fellow citizens and fellow citizens only.

Following this method, I will use the ideas of “harm,” “welfare,” and “autonomy” to organize the idea of political patienthood *per se* into the framework of Justice as Fairness. These ideas have precedents in the background culture of Western society, since they have been regularly invoked and discussed by liberals utilitarians, and many other moral and religious comprehensive doctrines, in both the academic and political realms. The concepts can be spelled out in a way that is sufficiently clear and simple, and the concepts can be shown to be “congenial” to our convictions by organizing them within the framework of the original position.

A working definition of “harm” and “welfare” can be drafted as follows. The welfare of an agent consists in the realization of her conception of the good, while the harm of an agent consists in the inhibition or restriction of the exercise of her moral powers.<sup>179</sup> The welfare of a patient might be understood as the

---

<sup>179</sup> Another approach to this working definition is to allow for cases where an agent receives his or her good but is nevertheless said to be harmed. It is more consistent with Rawls’s approach to limit an agent’s good to the realization of her conception of the good, and more consistent with the

satisfaction of basic preferences, such as the need for food, education, shelter, health, or clothing. However, the idea of satisfaction of basic preferences might not be the sort of idea that achieves pragmatic feasibility and widespread agreement in a liberal society. A more complicated but perhaps more feasible option involves neutrality with respect to the correct good-of-a-patient on behalf of representatives who represent political patients. These representatives would not know *which* theory correctly describes the good-of-a-patient, and would apply Rawls's maximin reasoning to attempt to secure the best outcome on the most pessimistic (yet still reasonable and minimally credible) normative theory that describes the good-of-a-patient. It might turn out that preference satisfaction is the correct good-of-a-patient, or pleasure, or some other conception of the good-of-a-patient, but the representative reasons using the conception that secures the best lot under the worst theoretical scenario. The preference-satisfaction approach has the virtue of simplicity, while the approach that is theoretically neutral with respect to the good-of-a-patient is likely more pragmatically feasible.

The specific set of basic preferences changes depending on what kind of non-agent patient is under consideration, but basic preferences in general are those preferences which must be satisfied in order for the patient to continue living a decent life, given the sort of life that patient can live. The harm that a patient can experience is defined as a frustration of these basic preferences, as in the case wherein the patient experiences cruelties such as torture, excessive pain

---

liberal tradition to steer clear of ideas that allow for paternalism towards agents. At the very least, the inclusion of cases where an agent is satisfied yet harmed requires a separate argument.



or punishment, deprivation of basic needs, or any other circumstances that prevent the patient from living a decent life.

A “decent life” is admittedly a vague notion, since there will inevitably be cases where it is either unclear what a decent life for an patient is because that patient cannot communicate his her preferences, or cases where the patient’s life is so impoverished due to some condition that a decent life is rendered impossible. In the first case, claims are made on *behalf* of the patient by a caregiver who has reasonable insight into the needs of the patient. This reasonable insight would consist in having the relevant background knowledge of medicine, nursing, veterinary medicine and so on. In the latter case, where a decent life is out of question, our political duties to these patients would be entirely *palliative*, in the sense that there would be duties to provide that patient an approximately decent life, or a decent life as far as possible, by means of comfort, care and assistance. With these exceptions aside, the idea of a “decent life” should be sufficiently clear when it is conceived as a restricted version of preference satisfaction, wherein the patient’s basic needs are fulfilled, and other preferences are satisfied within reasonable constraints.

A political conception of patienthood can be tentatively defined as the capacity to experience harm or welfare, and the capacity to experience a decent life. These are experienced phenomenally, and so are distinct from merely objective goods as in the case where “water is good for a plant” and “trajectory programming is good for a guided missile.” These capacities represent merely

sufficient conditions for patienthood since the two moral powers are equally sufficient for being owed duties of justice. Using Rawls's constructive approach, this political conception of patienthood can be worked into the original position by transforming it into reasonable constraints. In other words, the veil of ignorance, as an expression of the original position's reasonable constraints, can be modified such that representatives do not know the specific features of the member they represent, and only know that the member could be a non-agent patient or an agent. Thusly, the parties in the original position are ignorant of specific features of agents such as class, gender, age, conception of the good, and so on, and they are also ignorant of the specific features of patients such as species and level of cognitive or physical disability. The representatives would attempt to secure the abstract interests of the entity they represent in case that entity belongs to the least-advantaged members of society. Thus, the representative would agree to principles that secure basic rights and primary goods for the member if that member is an agent, and they would agree to principles that secure basic needs and a decent life for that member if that member is a patient. The latter principle would secure rights for patients such as a right to care, a right to comfort, or a right to rehabilitation in case of disability. The right to a caregiver and the right to care-giving facilities and institutions would ensure that the patient is secured a decent life. When a patient is a non-agent, these rights would translate into duties that institutions can fulfill.

One might argue that if a political conception of patienthood is implicit in the background culture of a Western democratic society, and we are able to

construct duties out of that conception, then the duties may nevertheless be of the wrong kind required for a political conception of justice. As discussed in an earlier section, duties of justice are duties that are legitimately coercively enforceable by the state. However, there are duties that a public can recognize and include in its background culture that are nevertheless not coercively enforceable. The conditions for depending when a duty is coercively enforceable were also described in an earlier section. The following is a clarification of the two conditions discussed earlier in order for a duty to be coercively enforceable:

- (1) The duty must be *practically feasible*, in the sense that we can expect that particular agents, after achieving wide reflective equilibrium, will recognize the duty and follow through with the expectations outlined by the duty.
- (2) The duty must also be publically recognized and widely held, in the sense that it is endorsed not merely by particular agents but by the *overlapping consensus* of reasonable comprehensive doctrines.

A coercively enforceable duty is entailed by the constructive procedure of the original position when it satisfies these conditions. The objection states that our duties to non-agent patients may be entailed by the constructive procedure (by the inclusion of a political conception of patienthood among the basic considered convictions), it may be practically feasible in that agents will agree to carry out

those duties, and there may even be consensus regarding the duty.<sup>180</sup>

Nevertheless, some duties will nevertheless fail to be coercively enforceable by the state. For example, it might be widely recognized that citizens have a duty not to cheat on their partners in monogamous relationships, but the state clearly has no place in punishing a citizen for failing to fulfill this duty. Another example of such a duty is the duty to recycle one's aluminum cans. There are financial rewards offered for recycling used cans, but one will not be fined or detained for disposing of cans in a dumpster. How can the political conception of justice distinguish cases where the three conditions are met and the duty is enforceable and cases where the three conditions are met but the duty is nevertheless unenforceable?

Earlier, I presented a list of exempting qualities that rendered a duty unenforceable by the state, even if it survives philosophical inquiry and achieves public recognition. These exemptions can be invoked to distinguish the two kinds of cases, and thereby establish that there are some duties to non-agent patients that

---

<sup>180</sup> A similar objection would state that there simply is no reasonable consensus regarding our duties to non-agent patients, because speciesism – the idea that the interests of certain species (viz. humans) have greater moral weight than the interests of other species – is incredibly prevalent. Hyper-Kantianism, the idea that *only* respect for autonomy matters in moral considerations, is another locus of possible dissent. There may also be religious views that value the sanctity of human life in a way that subordinates the interests of non-human animals. These views, so far as they subordinate the interests (or the good) of political non-agent patients, would have to be deemed *unreasonable* on my account, without arguing from some specific comprehensive doctrine. I would like to call into question the magnitude of the dissent afforded by these views. While extreme variants of these views will refuse to acknowledge duties to animals on *any* grounds, these variants seem unreasonable in that they are unwilling, in the process of reflective equilibrium, to consider the pragmatic benefits from within the context of their own doctrines to including the idea of political patienthood in political society. Like the Utilitarians who acknowledge the pragmatic value of liberal autonomy for the benefit of utility, the hyper-Kantians and speciesists could acknowledge the pragmatic value of the concept of political patienthood for the sake of Kantian autonomy or human interests. At least, the pragmatic emphasis in constructivism provides an avenue to mitigate dissent from speciesist views to acknowledge at least some *political* duties to non-agent patients.

are coercively enforceable. The first exemption is the case where the duties are only privately held, and stem exclusively from a reasonable comprehensive doctrine and not from the political conception of justice. This exemption is subsumed by the second condition which requires that duties find a public basis in overlapping consensus rather than merely a private basis in a reasonable comprehensive doctrine.

The second exemption involves cases of *trivial* duties. Small kindnesses, such as opening doors for strangers, fall under this category. Other examples can be found where contractual duties are present but enforceable only where they are not trivial. For example, citizens have a duty to pay the full price of a purchase, but if one comes up a penny short, and the merchant discovers this short-changing a day afterwards, it would be unreasonable for the law to intervene and force the citizen to scrounge up a penny and deliver it to the merchant. The case of recycling aluminum cans is also a trivial duty, because such recycling involves very little effort and cost on the part of a citizen. Cases of triviality are presumably identified by common sense, and trivial duties are not specified by the political conception of justice.

A third class of duties that are exempted from coercive enforcement are *supererogatory* duties, which are duties that are praiseworthy to perform but permissible to fail to perform. For example, it is praiseworthy to donate large amounts of one's wealth to charities, but no one is expected to provide such donations, and no one is blamed for failing to provide such donations. In Justice

as Fairness, these duties are not coercively enforceable because parties in the original position would not agree to their enforcement, since there is a risk that the coerced performance of these duties would bring an agent below the minimum degree of resources required for the exercise of their moral powers.

Duties that are *vague* or exceptionally *complicated* are also not coercively enforceable, because it would be practically infeasible for a citizen who agreed to these duties to carry them out. For example, the duty “be good to others” is too vague for a citizen to follow. An intractable duty might include excessive qualifications such as a duty assist fellow passengers on 747 planes in the event of an emergency while flying at least two kilometres above North Carolina on Saturdays. While this complex duty might turn out be one’s actual duty on some rare occasion, it should nevertheless stem from duties which are simple yet comprehensible, and not too simple as to be vague. The example of cheating on one’s partner seems to fall into this category, because

In “The Offense Principle,” Joel Feinberg suggests another class of moral duties that are unsuitable candidates for coercive enforcement – the merely offensive. Feinberg takes up Mill’s idea of the “harm principle,” which is the idea that the state’s enforcement of duties is legitimate just in case those duties prevent harm to one or more citizens.<sup>181</sup> He drafts a similar “offense principle” that states that the government can also step in when someone is “wrongfully offended” but

---

<sup>181</sup> Mill, “The Harm Principle,” 70.

not necessarily harmed.<sup>182</sup> Feinberg does not offer any hard and fast distinction between “mere offense” and “wrongful offense,” but he believes it has something to do with “personal privacy or autonomy.”<sup>183</sup> For Feinberg, privacy is “a privileged territory or domain in which an individual person has the exclusive authority of determining whether another may enter, and if so, when and for how long, and under what conditions.”<sup>184</sup> He does not mean merely that stepping inside a stranger’s “space bubble” might offend that stranger, although unwelcome experiences in close proximity account for a large number of offenses to privacy. Feinberg is also referring to cases where, without one’s consent, another learns or discloses one’s private information, such as cases of peeping or slandering. Depending on the extent of personal “territory,” this definition may also cover cases where one must unwillingly witness offensive conduct.

Justice as Fairness could incorporate an exemption of the merely offensive. Because there are rights associated with personal autonomy *and* harm in the modified principles of justice, there will be cases where citizens have their rights *qua* agents violated while not having their rights *qua* patients violated. These are the kinds of cases to which Feinberg is referring, and because they involve violations of basic rights, they can be coercively enforced. However, in cases where there is neither harm to a citizen *qua* patient or violation to a citizen *qua* agent, there is merely offense, and the state has no duty to prevent or punish these kinds of offenses.

---

<sup>182</sup> Feinberg, “The Offense Principle,” 85.

<sup>183</sup> Feinberg, “The Offense Principle,” 93

<sup>184</sup> Feinberg, “The Offense Principle,” 92.

A final class of exemptions are those cases where the duties would be *impractical* or *excessively costly* to enforce. This set of exemptions is partly covered by the practical feasibility condition, since there will be cases where citizens would fail to agree to duties which might be entailed by reflective equilibrium. For example, even if it turned out that vegetarianism was the morally correct diet, we could not reasonably expect reasonable and rational citizens to agree to vegetarianism, nor could we expect the state to enforce the diet and punish carnivores. There are also cases where it would be excessively difficult or costly to enforce a duty, such as the case of recycling cans. Punishing individuals for failing to recycle cans would be a greater burden on society than the worth of the recycled cans that the state recovers from unruly non-recyclers. So, these duties are exempted from the state's coercive enforcement.

There will inevitably be duties to non-agent patients that are exempted from the coercive enforcement by the state, because they meet one or more of these conditions. However, if a political conception of patienthood is adopted in the original position, then the law will be able to recognize violations of the rights of patients in cases of cruelty and mistreatment, and the state will have a duty to provide mental health services, education, care-givers, and so on to cases of severely cognitive disability. In the case of physical disability, there would be a state duty to provide the appropriate amount of assistance required for the citizen to make full use of his or her primary goods.



Because these duties are grounded on a publically recognized conception of patienthood with its respective moral capacities, these duties would not be exempted from enforcement on the basis that they derive from a merely private doctrine. Common sense should suffice to explain why duties to patients (such as providing care or preventing harm) are non-trivial. Parties in the original position would only agree to principles of justice that took violations of the rights of patients seriously, in the event that the entity they represent is merely a moral patient. On this basis, duties to patients are also not supererogatory, since the state can legitimately punish, say, acts of animal cruelty and can be sued if the appropriate services are not provided to disabled citizens.<sup>185</sup> We can also expect that some citizens will take offense to the fact that political patients are entitled to goods and services, since that would imply that the distributive scheme of the difference principle would grant agents a smaller share of wealth than they would have if political patients had no such entitlements. This offense, however, is unreasonable since it involves no violation of the rights of agents (since the exercise of their moral powers are already guaranteed by the first principle of justice). Offended citizens would have to take solace in the fact that the difference principle entitles agents to a greater share of wealth so far as it benefits the least-advantaged. As for vagueness or needless complexity, I have already argued that the moral capacities to experience harm and welfare are sufficiently

---

<sup>185</sup> Parties in the original position would require the state to be *accountable* for the rights and services to which a patient is entitled, so that the state has some incentive to ensure that the patient lives a decent life. If the parties in the original position merely acknowledged supererogatory duties to patients, they would thereby fail to secure the fundamental interests of the entities they represent.

clear when defined in terms of living a decent life, appropriate to the kind of life a given patient can live.

The most worrisome quality that might exempt duties to non-agent patients would be that these duties are impractical or costly to enforce. Some citizens might agree to perform duties to some kinds of political patients but not others, and not merely on some basis of bigotry, but because a citizen might not see in a given case how a patient is worthy of rights and entitlements. For example, in cases of psychopathy, we can expect reasonable people to experience moral outrage towards these patients instead of recognizing duties to provide mental health services and rehabilitation as far as possible. The basic structure could accommodate these cases over time by adjusting the education system to include relevant knowledge of political patienthood and the way that it figures into a well-ordered society. Other duties to non-agent patients, such as the state's provision of care and assistance, would cost the government significantly more resources than if these rights were not recognized, and the care of patients was left to the benevolence of agents. According to Rawls's circumstances of justice, however, such a society is founded in conditions of moderate scarcity, which means, while wealth is by no means infinite, there should be sufficient means to provide care for patients, perhaps at the expense of leisure and luxury to some agents. The state, then, would view duties to non-agent patients as having reasonable costs.

While the aforementioned list of cases where duties are exempted from being coercively enforced by the state may not be exhaustive, it should indicate that there are at least some duties to non-agent patients that warrant duties of justice towards them. The kinds of patients that political society would recognize can be worked out in detail on another occasion, but the conception of political patient as possessing the moral capacities for experiencing harm and welfare at the very least should account for the cases discussed earlier as “problems of extension.” The conception includes at least: those with severe cognitive or physical disability, or profound mental disorder, non-human animals capable of experiencing a decent life, Tay-Sachs children, and the “quite senile.” As discussed earlier, and confirmed by the conception of political patienthood, those in a persistent vegetative state or encephalic infants cannot be considered patients since they are incapable of a living any kind of a decent life.

By eliminating the Kantian requirement, and distinguishing between political agency and patienthood, Justice as Fairness should be able to respond to Nussbaum’s objection that it has simply no way of dealing with cases of severe disability. This method does not import any assumptions that Rawls’s constructivist approach does not already make; rather, it allows the approach to be truer to its constructive ideal by detaching itself from Kantian ideals while accounting for patients that the public recognizes as normatively important. The duties entailed by this conception may be either coercively enforceable because they meet the conditions of enforceability by being entailed by the original position, practically feasible and consensus-producing, or they may be exempted

from coercive enforceability if they are trite, too costly, etc. Either way, it is possible for Rawls's liberal state to recognize some duties to non-agent patients.

### **Conclusion**

In this thesis, I have examined Justice as Fairness to determine how “problems of extension” arise when the two moral powers are taken as necessary and sufficient conditions for any kind of political membership. I discussed the domain of justice as opposed to the domain of morality in order to determine what it would mean to “extend” a political conception of justice to cases outside of its professed applications, and concluded that a political conception of justice is extended to certain entities when it entails *coercively enforceable* duties to those entities. I explained the idea of Kantian constructivism in order to explain how Rawls thinks that his conception of moral personality – the two moral powers – figures into a political theory constructed out of our shared considered convictions as a Western democratic society. The two moral powers provide the elementary resources out of which Rawls's theory of justice is built, and so the principles of justice are designed to secure the interests of citizens insofar as they are reasonable and rational. I argued that the constructivist interpretation of Rawls fails to explain why the moral powers are *necessary* for being owed duties, and traced this idea back to Kant's moral philosophy, where the Kingdom of Ends is conceived in such a way as to preclude membership from entities that lack an autonomous will. The purview of the categorical imperative is limited to creatures with dignity, so that all moral patients are also moral agents. I argued

that it is best to remove this hyper-Kantian requirement from Justice as Fairness, because it is not necessary for constructing a political conception of justice, and threatens to undermine the resulting theory's politicality. This revision allows us to conceive of political patienthood as separate from political agency. I suggested that an intelligible and reasonable conception of political patienthood can be drafted using the ideas of harm and welfare, though the substantive content of the conception of political patienthood is best understood as whatever content is most likely to find consensus. I argued that there are coercively enforceable duties to political patients, and argued that these duties do not fall outside the domain of justice or find exemption by being merely private, trivial, supererogatory, vague, unduly complicated, merely offense-preventing, impractical or too costly. I conclude that eliminating the Kantian requirement allows Justice as Fairness to solve problems of extension that involve political patients, and that there are coercively enforceable duties to political non-agent patients.

## Bibliography

- Butcher, Mineka, and Hooley. *Abnormal Psychology*. Thirteenth Edition. 497.
- Elliot, Robert. "Rawlsian Justice and Non-Human Animals," *Journal of Applied Philosophy*, 1, pp. 95-106. 12. March 1984.
- Feinberg, Joel. "The Offense Principle." *Applied Social and Political Philosophy*. Ed. Elizabeth Smith and H. Gene Blocker. Prentice Hall: New Jersey 1994. P. 85.
- Henry Sidgwick, *The Methods of Ethics: Seventh Edition*, (Indianapolis: Hackett Publishing Company 1981).
- Hill, Thomas. "Kantian Constructivism in Ethics." *Ethics*. Vol. 99, No. 4 (Jul. 1989). The University of Chicago Press. pp. 752-770.
- Hobbes, Thomas. *Leviathan. Classics of Moral and Political Theory*. Ed. Michael L. Morgan. Hackett Publishing: Indianapolis, 2005. p. 608
- Hume, David. "Of Justice." *Applied Social and Political Philosophy*. Ed. Elizabeth Smith and H. Gene Blocker. Prentice Hall: New Jersey 1994. p 91.
- Kaufman, Alexander. "Kant and Constructivism." Paper presented at the annual meeting of the American Political Science Association, Marriott, Loews Philadelphia, and the Pennsylvania Convention Center, Philadelphia, PA, Aug 31, 2006 Online <PDF>. 2009- 02-02 <[http://www.allacademic.com/meta/p150482\\_index.html](http://www.allacademic.com/meta/p150482_index.html)>
- Kant, Immanuel. "Groundwork to the Metaphysics of Morals." *Practical Philosophy*. ed. M. Gregor. Cambridge University Press: Cambridge 1996.
- Kant, Immanuel. "Critique of Practical Reason." *Practical Philosophy*. ed. M. Gregor. Cambridge University Press: Cambridge 1996.
- Kant, Immanuel. "The Metaphysics of Morals." *Practical Philosophy*. ed. M. Gregor. Cambridge University Press: Cambridge 1996.
- Kant, Immanuel. "Towards Perpetual Peace." *Practical Philosophy*. ed. M. Gregor. Cambridge University Press: Cambridge 1996.
- Korsgaard, Christine. *Creating the Kingdom of Ends*. Cambridge University Press: Cambridge 1996.

- Korsgaard, Christine. *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*.
- Korsgaard, Christine. *The Sources of Normativity*. Oxford University Press: Oxford 2008.
- Levine, Andrew. "Rawls' Kantianism" *Social Theory and Practice*. Vol. 3 No. 1 (Spring 1974). pp. 47-64.
- Light, Andrew and Holmes Rolston, "Introduction: Ethics and Environmental Ethics," *Environmental Ethics: an anthology*, (Malden: Blackwell Publishers, 2003).
- Nozick, Robert. "The Importance of Liberty and Self-Ownership." *Applied Social and Political Philosophy*. Ed. Elizabeth Smith and H. Gene Blocker. Prentice Hall: New Jersey 1994. p 676.
- Nozick, Robert. "The Minimal State." *Applied Social and Political Philosophy*. Ed. Elizabeth Smith and H. Gene Blocker. Prentice Hall: New Jersey 1994. p 34.
- Nussbaum, Martha. "The Capabilities of People with Cognitive Disabilities," *Metaphilosophy*, 40(3-4), pp. 331-351. 21. July 2009.
- Nussbaum, Martha. *The Quality of Life*. Oxford University Press: Oxford 1993.
- Mill, John Stuart. "The Harm Principle." *Applied Social and Political Philosophy*. Ed. Elizabeth Smith and H. Gene Blocker. Prentice Hall: New Jersey 1994. p 70.
- O'Neill, Onora. "Constructivism in Rawls and Kant." *The Cambridge Companion to Rawls*. Ed. Freeman, Samuel (2003). pp. 347-367.
- O'Neill, Onora. "Constructivism vs. Contractarianism." *Ratio [0034-0006]*. Vol. 16. Iss. 4. (2003) pp. 319-331.
- O'Neill, Onora. "The Presidential Address: Constructivisms in Ethics" *Proceedings of the Aristotelian Society, New Series*. Blackwell Publishing on behalf of The Aristotelian Society. Vol. 89. (1988 - 1989), pp. 1-17.
- O'Neill, Onora. "Constructivism in Rawls and Kant." *The Cambridge Companion to Rawls*. Ed. Samuel Freeman. Cambridge University Press, 2002. Cambridge Collections Online. Cambridge University Press.

- Onora O'Neill, *Constructions of Reason*, (Cambridge: Cambridge University Press, 1989), 207.
- Pogge, Thomas W. "The Kantian Interpretation of Justice as Fairness". *Zeitschrift für philosophische Forschung*. Jan./Mar. Vol. 35. No. 1 Cambridge: Mass. 1981.
- Pritchard, Michael S, and Robison, Wade. "Justice and the Treatment of Animals: A Critique of Rawls," *Environmental Ethics*, 3, pp. 55-61, 7. Spring 1981.
- Rawls, John. "Justice as Fairness: Political not Metaphysical." *Philosophy and Public Affairs*. Vol. 14. No. 3. Pub. Blackwell Publishing (Summer, 1963) pp. 223-251.
- Rawls, John. "The Sense of Justice." *The Philosophical Review*. Vol. 72. No. 3. (Jul. 1963). Pub. Duke University Press on behalf of Philosophical Review. pp. 281-305.
- Rawls, *A Theory of Justice*, Cambridge: Belknap Press 1971.
- Rawls, John. "Kantian Constructivism in Moral Theory." *The Journal of Philosophy*. Vol. 77. No. 9 (Sep. 9, 1980), pp. 515-572. Journal of Philosophy, Inc. <<http://www.jstor.org/stable/2025790>> Accessed 8/12/2011.
- Rawls, *Justice as Fairness: A Restatement*, Ed. Erin Kelly, Cambridge: Harvard University Press, 2001.
- Rawls, *The Law of Peoples*, Harvard University Press, 2002.
- Rawls, *Political Liberalism*, Columbia University Press: New York, 2005
- Richardson, Henry S. "Rawlsian Social-Contract Theory and the Severely Disabled", *Journal of Ethics*. 10(4), pp. 419-462. 44. 2006.
- Robert, Taylor. "Rawls's Defense of the Priority of Liberty: A Kantian Reconstruction." *Philosophy and Public Affairs*. Vol. 31. No. 3. Blackwell Publishing (Summer, 2003) pp. 246-271.
- Scanlon, T.M. *What We Owe to Each Other*. Belknap Press: Cambridge 2000.
- Singer, Brent. "An Extension of Rawls' Theory of Justice to Environmental Ethics." *Environmental Ethics*, 10, pp. 217-232, 16, Fall 1988.
- Singer, Peter. *Practical Ethics*. Cambridge University Press: Cambridge 1993.



Thomas. "Rawls on Justice." *The Philosophical Review*. Vol. 82. No. 2. (Apr. 1973). Pub. Duke University Press on behalf of Philosophical Review. pp. 220-234.

VanDeVeer, Donald. "Of Beasts, Of Persons, and The Original Position," *The Monist*, 62, July 1979 pp. 368-377.