THE SIGNIFICANCE OF ROUSSEAU'S CONCEPT OF AMOUR-PROPRE IN RAWLS

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Xinghua Wang
May 2017

Copyright ©2017 by Xinghua Wang
All rights reserved

DEDICATION

To my Daughter, Julie Meinuo

ACKNOWLEDGEMENTS

I thank Dr. Reidy, who is my thesis supervisor, for his help with my dissertation throughout the past the two years. He reviewed every draft of each chapter of my dissertation and provided invaluable comments, without which, this dissertation would not be what it is now. I also thank him for lecturing and helping me in the several courses, in which I was a grader. I learned both how to improve my and the students' writings in these courses, and, more importantly, how to do philosophy by reading classical texts.

I also thank Dr. Garthoff for his help with my dissertation. He reviewed most drafts of my dissertation and provided invaluable suggestions. We met regularly to discuss these drafts and these discussions often give me inspiration.

I also thank Dr. Palmer and Dr. McAlpin for being members of the dissertation committee and giving constructive suggestions on some of the drafts of my dissertation. Dr. Palmer also helped me with my philosophical writings in the first two years of my PhD program, for which I am grateful.

I also thank all the other teachers who lectured me during the first three years of my PhD program. They showed me the exemplary way of being a teacher in one way or another. They include: Dr. Aquila, Dr. Berenstain, Dr. Boylu, Dr. Coffman, Dr. Cureton, Dr. Kohl, Dr. Nolt, and Dr. Shaw.

Last but not least, I thank all my friends for their company. Without them, these years wouldn't have been so enjoyable. I thank my parents and my mother-in-law for their selfless support during these years. I thank my daughter, Julie Meinuo, for bringing so much joy to my life. I thank my husband Jun for proofreading almost every draft of my dissertation, and for his love and being the best friend of mine.

ABSTRACT

This dissertation defends the view that there is a Rousseauvian interpretation of Rawls's political philosophy by focusing on the significance of amour-propre in Rawls's political philosophy. In the first chapter, I introduce my central thesis and chapter arrangements and compare my Rousseauvian interpretation with other interpretations of Rawls. In the second chapter, I introduce Rousseau's concept of amour-propre and try to defend Rawls's wide view of amour-propre, according to which, amour-propre has both a positive and a negative form. In the third chapter, I argue that Rousseau's concept of amour-propre plays a significant role in Rawls's conception of justice as fairness. Thus, I show that one of the main reasons why parties in the original position would choose Rawls's two principles of justice over other conceptions of justice is that justice as fairness meets the demands of amour-propre while other conceptions of justice do not. In the fourth chapter, I argue that Rousseau's concept of amour-propre plays a significant role in Rawls's stability argument. This argument has three parts. The first part involves showing that Rawls's moral psychology is an illustration of Rousseau's thesis that moral sentiments are derived from natural sentiments. The second part argues that Rawls's congruence argument and his overlapping consensus argument are based on Rousseau's conception of persons. The third part argues that Rawls's argument from the absence of special psychology is also grounded in Rousseau's concept of amour-propre. In the fifth chapter, I argue that Rousseau's concept of amour-propre plays a significant role in Rawls's law of peoples by showing that three themes of Rawls's law of peoples, the two international original positions, the idea of a realistic utopia and the distinction between peoples and states, parallel, and are grounded in, Rousseau's concept of amour-propre. In the final chapter, I consider two objections against my assumption that the legitimate social bases of self-respect are equal basic rights and

liberties and their substantive fair value. I argue that these objections are groundless. I thus conclude that there is a Rousseauvian interpretation of Rawls's justice as fairness and law of peoples.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1

1. The Central Thesis of This Dissertation and The Aim of This Chapter 1

2. The Kantian Interpretation and Its Shortcomings 2

2.1 The Kantian Interpretation of Rawls 3

2.2 The Objections Against the Kantian Interpretation 5

2.3 The Existing Defenses of Rawls's Kantian Interpretation 7

2.4 Understanding Rawls's Kantian Interpretation from a Rousseauvian Perspective 13

3. The Rousseauvian Interpretation of Rawls 17

3.1 Other Roussseauvian Interpretations of Rawls and Their Shortcomings 18

3.2 My Rousseauvian Interpretation of Rawls 20

4. The Task of Each Chapter, and How It Contributes to the Overall Project 24

CHAPTER 2 ROUSSEAU'S CONCEPT OF AMOUR-PROPRE 29

1. The Traditional View of Amour-propre 31

2. The Wide View of Amour-propre 34

3. An Argument Against the Wide View of Amour-propre 40

4. My Defense of the Wide View of Amour-propre 43

4.1 The Role of Pity vs the Role of Self-love or Self-esteem in the Second *Discourse* 43

4.2 The Role of Pity vs the Role of Self-love or Self-esteem in *Emile* 45

4.3 The Role of Pity vs the Role of Self-love or Self-esteem in Rousseau's Other Works 50

5. Conclusion and Some Afterthoughts 53

CHAPTER 3 THE SIGNIFICANCE OF ROUSSEAU'S CONCEPT OF AMOUR-PROPRE IN RAWLS'S JUSTICE AS FAIRNESS 57

1. A Brief Introduction to Justice as Fairness 59

2. The Argument from Self-respect for the Two Principles of Justice 62

2.1 The Utilitarian Principle Fails to Meet the Demand of Amour-propre.....64

2.2 The Two Principles of Justice Meet the Demand of Amour-propre.....70

3. Conclusion 83

CHAPTER 4 THE SIGNIFICANCE OF ROUSSEAU'S CONCEPT OF AMOUR-PROPRE IN RAWLS'S STABILITY ARGUMENT..... 85

1. Amour-propre and Rawls's Moral Psychology 87

1.1 A Brief Introduction to Rawls's Moral Psychology 88

1.2 The Parallel Relations Between Rousseau's Concept of Amour-propre and Rawls's Moral Psychology 94

2. Amour-propre, the Congruence Argument and the Overlapping Consensus Argument 106

2.1 Amour-propre and the Congruence Argument in *Theory*.....107

2.2 Amour-propre and the Overlapping Consensus Argument in *Political Liberalism* 110

3. Amour-propre and Rawls's Argument from the Absence of Special psychology ...113

CHAPTER 5 THE SIGNIFICANCE OF ROUSSEAU'S CONCEPT OF AMOUR-PROPRE IN RAWLS'S LAW OF PEOPLES..... 117

1. Amour-propre and Rawls's Law of Peoples 119

2. Rawls's Law of Peoples Meets the Demand of Amour-propre.....124

2.1 Why Does Rawls's Law of Peoples Meet the Demand of Amour-propre? 125

2.2 Amour-propre and Stability in the Law of Peoples 129

2.3 Amour-propre and Rawls's Arguments against Political Realism and Distributive Cosmopolitanism.....132

3. Objections and Responses.....138

3.1 The First Objection 139

3.2 The Second Objection.....144

4. Conclusion 148

CHAPTER 6 CONCLUDING REMARKS..... 150

1. Chinese Nationalists and Self-respect 151

1.1 Chinese Nationalists' View on Self-respect.....152

1.2 Why Can the Tradition of Old China not be the Proper Bases of Self-respect?.....154

2. Luck Egalitarianism's Objections against Rawls 160

2.1 Joshua Cohen's Argument from Self-respect for Rawls's Difference Principle 161

2.2 G. A Cohen's Objection Against Rawls's Difference Principle.....161

2.3 My Argument Against G. A Cohen's Objection.....164

3. Conclusion 168

BIBLIOGRAPHY..... 169

VITA 183

ABBREVIATIONS

REFERENCES TO WORKS OF ROUSSEAU

E *Emile*, or *On Education*, trans. Allan Bloom (New York: Basic Books, 1979).

SC *The Social Contract*, in *The Social Contract and other Later Political Writings*, trans. Victor Gourevitch (Cambridge: Cambridge University Press, 1997), 39-152.

DI Discourses on Inequality (*Discourse on the Origin and Foundations of Inequality among Men*), in *The Discourses and Other Early Political Writings*, trans. Victor Gourevitch (Cambridge University Press, 1997), 111-222.

PE *Discourse on Political Economy*, in *The Social Contract and Other Later Political Writings*, trans. Victor Gourevitch (Cambridge University Press, 1997), 3-38.

RJJ *Rousseau, Judge of Jean-Jacques: Dialogues*, trans. Judith R. Bush, Christopher Kelly, and Roger D. Masters, in *The Collected Writings of Rousseau* (Hanover, NH: University Press of New England, 2001), vol. I.

OW *The Collective Writings of Jean-Jacques Rousseau*, trans. Judith R. Bush, Christopher Kelly, and Roger D. Masters (Hanover, NH: University Press of New England, 2001).

REFERENCES TO WORKS OF JOHN RAWLS

TJ *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971).

PL *Political Liberalism* (New York: Columbia University Press, 1993).

LP *The Law of Peoples* (Cambridge, Mass.: Harvard University Press, 1999).

JF *Justice as Fairness* (Cambridge, Mass.: Harvard University Press, 2001).

LHPP *Lectures on the History of Political Philosophy*, ed. Samuel Freeman (Cambridge, Mass.: Harvard University Press, 2007).

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

CP *Collected Papers*, ed. Samuel Freeman (Cambridge, Mass.: Harvard University Press, 1999).

CHAPTER 1 INTRODUCTION

1. The Central Thesis of This Dissertation and The Aim of This Chapter

The central thesis of my dissertation is that there is a Rousseauvian interpretation of Rawls's political philosophy. I will defend this thesis by focusing on the significance of Rousseau's concept of amour-propre in Rawls's justice as fairness, stability argument and the law of peoples in the following chapters.

This chapter aims to provide the general picture of my interpretation of Rawls and compare it with other interpretations of Rawls. I will first introduce the Kantian interpretation of Rawls. I then consider two difficulties facing the Kantian interpretation. The first difficulty is that the Kantian interpretation fails to explain why for Rawls acting from the two principles of justice is acting autonomously. I argue that this difficulty can be resolved by my Rousseauvian interpretation because Rawls's justice as fairness is grounded in Rousseau's (and so also Kant's) idea of autonomy. The second difficulty is that the Kantian interpretation fails to emphasize the importance of stability in Rawls. I argue that this difficulty can also be resolved by my Rousseauvian interpretation because Rousseau's concept of amour-propre (correctly interpreted by Kant) plays a significant role in Rawlsian stability. Next, I will briefly introduce several themes of the Rousseauvian interpretation of Rawls. I argue that some difficulties facing these themes can be resolved by my Rousseauvian interpretation. First, I argue that Freeman's Rousseauvian interpretation does not explain how a person who is self-interested is capable of the sense of justice for both Rousseau and Rawls. I argue that my interpretation explains this difficulty — for both Rousseau and Rawls, the moral sentiments of the sense of justice is originated from the natural sentiments of self-love or self-respect. Second, I argue that

Bercuson's Rousseauvian interpretation mainly focuses on the common aspects of the Hegelian and the Rousseauvian interpretations of Rawls, and it fails to emphasize the importance of amour-propre in Rawls's early and later works. My interpretation addresses this difficulty because it focuses on the significance of amour-propre in Rawls. Third, I argue that Neuhouser's Rousseuvian interpretation does not provide an adequate explanation for why Rousseau's idea of the general will parallels Rawls's idea of public reason. I argue that my interpretation explains away this difficulty because amour-propre plays a significant role in both the formation of the general will and that of public reason.

2. The Kantian Interpretation and Its Shortcomings[1]

In this section, I will first introduce Rawls's Kantian interpretation of justice as fairness. Next, I will introduce the objections against the Kantian interpretation. After introducing these thoughts, I will examine the existing defenses of the Kantian interpretation. I argue that these defenses fail to explain why for Rawls acting from the two principles of justice is acting autonomously. I argue that my Rousseauvian interpretation provides such an explanation. Finally, I consider the objection that the Kantian interpretation fails to emphasize the importance of stability in Rawls. I argue that my Rousseauvian interpretation, which focuses on Rousseau's concept of amour-propre, plays a significant role in Rawlsian stability.

[1] There are other interpretations of Rawls, such as Andres De Francisco (2006) who puts forward a republican interpretation of Rawls's political liberalism and Jeffrey Bercuson (2014) who advances a Hegelian interpretation of Rawls's justice as fairness. Francisco's interpretation focuses on the importance of citizenship in Rawls's political philosophy. He interprets Rawls's reasonableness as civic virtue which underlies the stability of the well-ordered society. Since he appeals to Rousseau's idea of the general will and his idea of political liberty in support of his interpretation, his republican interpretation and my interpretation are mutually supportive. Bercuson argues that Rawls is indebted to Hegel's concept of reconciliation, in particular, Hegel's claim that our human nature is reconcilable to a legitimate and stable political institution. As we will see later, Rawls is also indebted to Rousseau in the same way. I argue that Bercuson's interpretation mainly focuses on the common aspects of the Hegelian and the Rousseauvian interpretations of Rawls and so it fails to emphasize the importance of amour-propre in Rawls.

2.1 The Kantian Interpretation of Rawls

The Kantian interpretation is first endorsed by Rawls himself in *A Theory of Justice*, and latter supported by others such as Stephen Darwall (1976; 1980), Bernard H. Baumrin (1976), Paul Guyer (2000), Nicholas Tampio (2007), and Robert Taylor (2003; 2011).[2] Rawls himself titiles Section 40 of *A Theory of Justice*, "The Kantian Interpretation of Justice as Fairness." Rawls argues that his interpretation is based on Kant's notion of autonomy. He writes:

> For one thing, he [Kant] begins with the idea that moral principles are the object of rational choice. They define the moral law that men can rationally will to govern their conduct in an ethical commonwealth. Moral philosophy becomes the study of the conception and outcome of a suitably defined rational decision. This idea has immediate consequences. For once we think of moral principles as legislation for a kingdom of ends, it is clear that these principles must not only be acceptable to all but public at well. Finally Kant supposes that this moral legislation is to be agreed to under conditions that characterize men as free and equal rational beings. The description of the original position is an attempt to interpret this conception. (TJ 221)

In his original position argument, Rawls argues that parties that are free and equal rational beings under the veil of ignorance would choose the principles of justice over other conceptions of justice. Now Rawls argues that the premises of this argument parallel Kant's idea of the social

---

[2] Paul Guyer's Kantian interpretation of Rawls, differs from the interpretation others in that it focuses on Kant's political philosophy, especially his conception of right. He argues that "Rawls's two principles of justice are the necessary reconstruction of Kant's own conception of right" (Guyer, 2000, p.285). Nicholas Tampio (2007) also has a different Kantian interpretation of Rawls, he argues that Rawls's political liberalism is indebted to Kant's idea of autonomy and that Rawls's idea of public reason parallels Kant's idea of public reason in "What is Enlightenment?" Richard W. Miller (2012) also has a different Kantian interpretation of Rawls. He argues that Rawls's law of peoples "was modeled on Kant's claim that international justice consists of terms of confederation suitable to free peoples, terms largely concerned with the keeping of peace" (p.297).

contract and his idea of moral persons. First, Rawls argues that he and Kant share a contractarian assumption that moral principles must be public and acceptable for all. Second, he argues that they agree that a moral theory must characterize moral agents as free and equal rational beings. In particular, he argues that his idea of autonomy parallels Kant's idea of autonomy:

> Kant held, I believe that a person is acting autonomously when the principles of his action are chosen by him as the most adequate possible expression of his nature as a free and equal rational being……Now the veil of ignorance deprives the persons in the original position of the knowledge that would enable them to choose heteronomous principles. (TJ 222)

Here Rawls argues that the veil of ignorance ensures that the principles of justice chosen in the original position are not heteronomous, and so citizens who act on, and from, the principles of justice are not heteronomous. In other words, the people who are represented by parties in the original position are autonomous because they act on and from the principles of justice. Rawls continues:

> The principles of justice are also analogous to categorical imperatives. For by a categorical imperative Kant understands a principle of conduct that applies to a person in virtue of his nature as a free and equal rational being. The validity of the principle does not presuppose that one has a particular desire or aim. Whereas a hypothetical imperative by contrast does assume this: it directs us to take certain steps as effective means to achieve a special end…The argument for the two principles of justice does not assume that the parties have particular ends, but only that they desire certain primary goods. These are things that it is rational to want whatever else one wants. (TJ 223)

Rawls argues that since the principles of justice are chosen by free and equal rational agents who do not have particular ends, these principles are categorical. He assumes that parties in the original position are rational, that is, they desire certain primary goods. He argues that moral agents are not heteronomous simply because they desire certain primary goods; for the preference for primary goods is derived "from only the most general assumption about rationality and the conditions of human life" (TJ 223).

To sum up, Rawls argues that there is a Kantian interpretation of justice as fairness for three reasons. First, he and Kant share a conception of autonomy. A person is autonomous if he or she acts on the principles of justice chosen in the original position by free and equal rational beings. Second, his principles of justice parallel Kant's categorical imperative. Since the principles of justice are chosen by free and equal rational beings in the original position, they are not hypothetical. The principles of justice express the free and equal rational nature of the parties in the original position, and so they do not presuppose a special end or a particular desire. Third, he and Kant share a conception of rationality. Rationality is not instrumental in the sense that it is not purely used to pursue a particular desire or a special end. Rather, it is used to pursue the common interests.

2.2 The Objections Against the Kantian Interpretation

Many commentators either argue that Rawls misunderstands Kant, or argue that some aspects of Rawls's justice as fairness are not Kantian.[3] Some of these commentators argue that

[3] See, for example, Oliver Johnson (1974), Andrew Levine (1974), Allan Bloom (1974), H. E. Mason (1976), Agneta Sutton (1979), Thomas Pogge (1981), Michael Sandel (1997), Bonnie Honig (1993), Onora O'Neill (1989), Larry Krasnoff (1999), Frederick Neuhouser (2008) and Jeffrey Bercuson (2014).

the individuals in Rawls's original position are not Kant's noumenal self.[4] Others argue that since political and social institutions play a significant role in shaping the moral character of the individuals in Rawls's well-ordered society of justice as fairness, these individuals are heteronomous for Kant.[5] Some of these commentators suggest that had Rawls interpreted Kant correctly, he would have found that his overall project in *A Theory of Justice* is inconsistent with Kant's notion of autonomy.[6] Most of them agree that Rawls is mistaken in believing that he and Kant share the common ideas of autonomy, categorical imperative and rationality.[7]

First, Rawls's critics argue that Rawls and Kant do not share a common idea of autonomy.[8] They argue that, for Kant if a person is autonomous, the motive for his action can only be respect for the moral law. And so if wants and inclinations play any role in motivating a moral agent, that moral agent is not autonomous. They argue that Rawls seems to be using autonomy in a quite different way: for Rawls, a person is autonomous even though that person is motivated by the desire to promote his own interest. Thus, they argue that there is no parallel relation between Rawls's and Kant's ideas of autonomy.

Second, Rawls's critics argue that Rawls's principles of justice are not analogous to Kant's categorical imperative. On the one hand, they argue that Rawls's principles of justice, chosen by free and equal rational agents, are not categorical imperatives.[9] On the other hand,

---

[4] For example, see Andrew Levine (1974, p.57), Michael Sandel (1997, p.13), and Frederick Neuhouser (2008, p.258).
[5] See Jeffrey Bercuson (2014, pp.11-2).
[6] For example, see H. E Mason (1976, p.52-4).
[7] See Oliver Johnson (1974;1977), Andrew Levine (1974), Allan Bloom (1974), H. E. Mason (1976), Agneta Sutton (1979), and Thomas Pogge (1981).
[8] See Oliver Johnson (1974, pp.60-2;1977, p.251), Andrew Levine (1974, pp.48-54), Allan Bloom (1974, p.656), H. E. Mason (1976, p.50), Agneta Sutton (1979, p.138), Thomas Pogge (1981, pp.48-9) and Jeffrey Bercuson (2014, pp.11-2).
[9] See Oliver Johnson (1974, pp.62-3), Andrew Levine (1974, pp.54-5), Allan Bloom (1974, pp.256-7), H. E. Mason (1976, p.50), Agneta Sutton (1979, p.138), and Thomas Pogge (1981, pp.47-8).

they argue that Kant's categorical imperative procedure is a procedure made by actual individuals in the world to produce possible agreement to serve as a negative check on the maxims that individuals bring about and that it is not, as Rawls argues, a procedure made by hypothetical individuals to produce hypothetical agreement to produce moral principles.[10] Thus, they argue that there is no parallel between Rawls's principles of justice and Kant's idea of the categorical imperative and between Rawls's original position procedure with Kant's categorical imperative procedure.

Third, Rawls's critics argue that Rawls and Kant do not share the same idea of rationality.[11] They argue that the parties in Rawls's original position are assumed to be only capable of means-ends reasoning. However, for Kant, the "true vocation of reason" is "to produce a will that is good, not perhaps as a means to other purposes, but *good in itself*, for which reason was absolutely necessary" (Kant, 1996, p.52). Thus they argue that Rawls's idea of rationality in the original position does not parallel Kant's idea of reason.

2.3 The Existing Defenses of Rawls's Kantian Interpretation

Other commentators try to defend Rawls's interpretation of Kant.[12] First, some of these commentators argue that Rawls's interpretation of Kant's notion of autonomy is not mistaken.[13] They argue that for Kant, autonomy is to act on the law that one gives to oneself. Rawls's two

[10] See Onora O'Neill (1989, p.217) and Larry Krasnoff (1999, p.400-1).
[11] See Oliver Johnson (1974, pp.63-6), Allan Bloom (1974, p.256), H. E. Mason (1976, p.50), Agneta Sutton (1979, p.136), and Thomas Pogge (1981, p.48).
[12] See, for example, Stephan Darwall (1976; 1980), Bernard H. Baumrin (1976), Paul Guyer (2000) and Robert Taylor (2003; 2011), Nicholas Tampio (2007) and Alexander Kaufman (2012).
[13] See Stephen Darwall (1976, pp.166-7), Bernard H. Baumrin (1976, p.56).

principles of justice are chosen in the original position by all free and equal rational beings. And so to act on and from the principles of justice is to act autonomously.

Second, some of these commentators argue that Rawls's principles of justice parallel Kant's categorical imperative.[14] Darwall (1976, p.167) argues that since the principles of justice are chosen in the original position by all free and equal rational beings, they are practical laws that require that "one be capable of regarding it as a principle which could be willed by all other rational beings also." Guyer (2000) argues that since primary goods are necessary for any rational ends, they are "means that are recommended by pure reason itself" (p.271). Taylor (2011) argues that social primary goods are not merely an instrument for the realization of our rational ends, but also necessary for the exercise and development of our two moral powers (cf. p.37).

Finally, some of these commentators argue that Rawls's idea of rationality is close to Kant's idea of reason. For example, Darwall (1976, p.168) argues that rational agency itself requires the possession of the primary goods. Baumrin (1976, p.57) argues that "In the original position prudence for me is prudence for humanity, and prudence for humanity is just the same as legislating for a kingdom of ends." Taylor argues that Rawls's idea that individuals possess the capability to form, revise and pursue their own conception of the good, is close to Kant's idea of prudential reasoning. (cf. Taylor, 2011, p.25).

But Rawls's critics might argue that Rawls's Kantian interpretation cannot be defended because it is simply not Kant's view that an autonomous person is merely a person who acts from a self-legislating law. They might argue that acting from a self-legislating law is only a

[14] Even though Kaufman (2012, p.13) argues for a Kantian interpretation of Rawls's justice as fairness, he does not argue for the parallel relation between Rawls's original position and Kant's categorical imperative procedure, but for the parallel relation between Rawls's original position and Kant's social contract theory.

necessary condition, but not a sufficient condition for autonomy. For example, Oliver Johnson (1977) says:

> Suppose an individual adopted the following rule of action for himself: Act always in such a way as to keep your reputation. In fulfillment of this rule he always tells the truth and never lies. According to Baumrin's [and Rawls's] interpretation of Kant's notion of autonomy, the individual in question would be acting autonomously, for his action would be pursuant to a rule, hence in accordance with law, and his rule would be self-legislated. He would be fully responsible for his acts, for they would be done in fulfillment of a self-imposed rule. Yet, if we look at Kant, we find him using the same illustration as I have — but as a case of *heteronomous* action. (Oliver Johnson, 1977, p.278)

Kant does say that hypothetical imperatives take the form that "I should do something because I will something else" and that the categorical imperatives take the form that "I ought to act in such or such a way even though I have not willed anything else" (Kant, 1996, p.89). And he says that "the former [the hypothetical] says that I ought not to lie if I will to keep my reputation" (ibid). But it seems to me that for Kant acting from a self-legislating law is both necessary and sufficient for autonomy because Kant did not think the person who always tells the truth in order to keep his reputation is acting from a self-legislating law. As Kant says:

> If the will seeks the law that is to determine it *anywhere else* than in the fitness of its maxims for its own giving of universal law — consequently if, in going beyond itself, it seeks this law in a property of any of its objects — *heteronomy* always results. The will in that case does not give itself the law; instead the object, by means of its relation to the will, gives the law to it. This relation,

> whether it rests upon inclination or upon representations of reason, lets only hypothetical imperatives become possible: I ought to do something *because I will something else*. (ibid.)

While Kant suggests that acting from the law one gives oneself is acting autonomously, He also seems to suggests that a person who is motivated by his self-interest is not acting autonomously; for he is not acting from pure practical reason alone. And so for Kant, it seems that acting from the two principles of justice which are chosen in the original position by free and equal rational agents is not acting autonomously; for the two principles of justice are not chosen by pure practical reason alone. Rather they are chosen by the parties who are not reasonable in the original position. While the parties represent reasonable citizens in the well-ordered society of justice as fairness, they are not reasonable themselves. The parties are assumed to be free and equal rational beings. They would choose the two principles of justice in the original position not because these principles express their nature as reasonable but because these principles express their nature as free and equal rational beings. In other words, the two principles of justice are chosen because they express their self-love nature. But Kant insists in a number of places that the principle of self-love cannot be willed as a universal law (cf. ibid., p.44; p.70; pp.158-60).[15] For example, he says:

> But suppose that finite rational beings were thoroughly agreed with respect to what they had to take as objects of their feelings of pleasure and pain and even with respect to the means they must use to obtain the first and avoid the other; even then they could by no means pass off *the principle of self-love* as a *practical law*; for, this unanimity itself would still be only contingent. (ibid.,

[15] This is pointed out by Kerstin Budde (2007, p.350).

p.159-60)

Kant thus has explicitly said that the principle of self-love cannot be a practical law despite the fact that it is thoroughly agreed by finite rational beings. Even though it is true that every person desires pleasure and tries to avoid pain, this maxim cannot be willed as a universal practical law because it has a subjective determining ground, an empirical foundation. He says: "The determining ground would still be only subjectively valid and merely empirical and would not have that necessity which is thought in every law, namely objective necessity from a priori grounds" (ibid., p.160). Since the desire for the primary goods also has an empirical foundation, it does not have the "objective necessity from a priori grounds". Thus, it is clear that for Kant acting from the two principles of justice is not acting from pure practical reason alone, but from empirical practical reason. And so it is puzzling why for Rawls acting from the two principles of justice is acting autonomously in the Kantian sense.

Rawls himself says in *A Theory of Justice* that he does "not wish to argue here for this interpretation on the basis of Kant's text" (TJ 221). He later says in his *Lecture* that "in presenting Kant's moral philosophy, I have played down the role of the a priori and the formal" (LHMP 275). He also criticizes Kant's categorical imperative procedure because Kant's reasoning rejects any maxim of mutual aid (LHMP 172-3).[16] He acknowledges that his Kantian interpretation makes two revisions to Kant's ethical theory. The first revision is to "give more content to the will of ideal agents in deciding whether they can will an adjusted social world"

[16] Rawls argues that the duty of mutual aid cannot be willed as a universal law for Kant because it requires us to help others in circumstances in which we may not want to help. Kerstin Budde (2007, p.346) argues that Rawls's reading of Kant's CI-procedure is superficial and misleading. He argues that the maxim of mutual aid cannot be willed as a universal law only when the maxim is expressed from self-love. He argues that since a possible maxim of mutual aid can be expressed as "If I see someone in need and I can help, I will, because I will to treat human beings as ends in themselves," it contains no contradiction of the will. Budde fails to see that for Rawls (and for Rousseau) self-love is compatible with the maxim of mutual aid and so it is puzzling for him why Rawls insists the maxim of mutual aid must be expressed from self-love.

(LHMP 173). This content is the true human needs (or the desires for social primary goods). The second revision is "to specify further the point of view from which these decisions about social worlds are made" (LHMP 173). But one may wonder to what extent Rawls's Kantian interpretation, which makes these revisions, is still Kantian. For example, Budde (2007) argues that Rawls is not a Kantian but that he gives a Rawlsian interpretation to Kant. He argues that Rawls's interpretation of Kant, which emphasizes the role of true human needs but neglects the role of a priori and the formal, has deviated from "the very essence of the natural of [Kant's] moral theory" (p.354). He argues:

> Rawls's rejection of the capacity of pure reason alone to ground moral principles through a priori and formal reasoning pushes him to an attempt to derive objective moral principles partly through empirical reasoning. Well aware that, as such, empirical reasoning is based on desires which are different for each person, Rawls tries to argue for universal human needs, the 'rational wanting' of which can then explain the force and authority of moral principles. However, this justification remains conditional: the moral principles hold as long as one desires the true human needs that the principles guarantee. (ibid.)

I agree with Buddle that Rawls rejects Kant's idea of pure practical reason that Rawls's Kantian interpretation has an empirical foundation, universal human needs, which appear to be absent in Kant's practical reasoning. But I think Buddle is mistaken to say that Rawls's interpretation of Kant has deviated the essence of the nature of Kant's ethics. As we will see, Rawls's interpretation of Kant can be understood from a Rousseauvian perspective.

2.4 Understanding Rawls's Kantian Interpretation from a Rousseauvian Perspective

As Rawls acknowledges, his Kantian interpretation is not strictly based on Kant's text. As we have seen, it is not clear why Rawls's justice as fairness is grounded in Kant's idea of autonomy. In particular, it is not clear why for Rawls acting from Rawls's two principles of justice is acting autonomously in the Kantian sense. Rawls's proponents fail to give us a persuasive response to the objections against the Kantian interpretation. I think that the key to understanding Rawls's Kantian interpretation is to see it from a Rousseauvian perspective. As Rawls writes:

> Those who think of Kant's moral doctrine as one of law and guilt badly misunderstand him. Kant's main aim is to deepen and to justify Rousseau's idea that liberty is acting in accordance with a law that we give to ourselves. And this leads not to a morality of austere command but to an ethic of mutual respect and self-esteem. (TJ 225)

Rawls explicitly says here that his idea of autonomy is both Kantian and Rousseauvian. For Rawls, both Kant and Rousseau maintain that autonomy is acting from the law one gives oneself. Thus, to understand why Rawls insists that acting from the two principles of justice is acting autonomously, we need to see Rawls's Kantian interpretation from a Rousseauvian perspective. Rousseau says that by subjecting oneself to the general will (the law), one is more free than one was in the state of nature because in the state of nature one was driven by appetites, but in the civil state, one's appetites are restrained by reason (SC 53). Rawls's original position argument is grounded in Rousseau's idea of autonomy. Rawls assumes that parties in the original position are free and equal rational beings. This assumption about individuals in the original position parallels Rousseau's assumption about individuals in the state of nature. Rawls claims that acting

from the two principles of justice chosen in the original position by free and equal rational beings is acting autonomously. This claim also parallels Rousseau's view that subjecting oneself to the general will (the law) makes one more free than one was in the state of nature. The reason why Rawls thinks that the two principles of justice are categorical, despite the fact that they are chosen by rational agents who all desire the social primary goods, is that these primary goods are not a particular person's private interest but the common interest of all. Rousseau distinguishes the will of all from the general will: "the latter looks only to the common interest, the former looks to private interest, and is nothing but a sum of particular wills" (SC 60). Thus, the general will is not the will deprived of all personal interest, but the will deprived of all private interest. Since Rawls's rational agents are assumed to have only the common interest of exercising and developing their two moral powers, they are deprived of all private interest. And so these rational agents are not heteronomous in the Rousseauvian sense. Thus, if, as Rawls suggests, Kant's main aim is to deepen and justify Rousseau's idea of autonomy, Rawls's rational agents are not heteronomous for Kant. It also follows that from a Rousseauvian Kantian perspective, acting from Rawls's two principles of justice is acting autonomously. Thus, it seems to me, that the reason why Rawls thinks that there is a Kantian interpretation of his justice as fairness is that he reads Kant in a Rousseauvian way. As Rawls writes in his *Lecture*: "it was not until I connected the *Second Discourse* with Kant's remarks here that I felt I finally understood what either of them was saying" (LHPP 200). Kant's interpretation of Rousseau's concept of amour-propre is endorsed by Rawls. Based on this interpretation, Rawls supposes that it is in a Kantian spirit to include self-love in practical reasoning. He assumes that Kant appreciates Rousseau's concept of amour-propre, especially the reciprocal feature of amour-propre, according to which, a person who has a proper sense of amour-propre is ready to grant equal standing to others. Thus, even

though the parties in the original position are assumed to be rational, not reasonable, since they would choose the principles of justice that are publically acceptable for others, the two principles of justice are reasonable. Thus, from a Rousseauvian Kantian perspective, acting from the two principles of justice are acting autonomously.

It is sometimes argued that the Kantian interpretation of Rawls fails to emphasize the importance of stability. For example, Jon Garthoff (2013) argues that Taylor's Kantian interpretation "fails to appreciate how deeply this orientation [stability] figures in Rawls' thinking" (p.285). He argues that Taylor's failure to appreciate the importance of stability in Rawls leads him to accuse Rawls of making a number of mistakes that Rawls isn't committed to. Jeffrey Bercuson (2014) also argues that Rawlsian stability is not Kantian.[17] He writes: "Rawls is unconvinced by the supposed natural immediacy of the moral law, and when Rawls himself acknowledges the socializing or pedagogical function of group practices values and (most importantly) institutions, he sees this as turning away from Kant" (Bercuson, 2014, p.12).

The importance of stability in Rawls's works has been emphasized by many. For example, Jon Garthoff (2016) emphasizes the importance of stability in Rawls's justice as fairness.[18] He argues that one of the main reasons why justice as fairness is more defensible than other conceptions of justice is that justice as fairness is more stable than others. He further argues that the criterion of publicity and that of reciprocity are constituents of stability, and that since justice as fairness meets these two criteria, justice as fairness is more stable than other conceptions of justice.

---

[17] See Todd Hedrick (2015, p.296).
[18] Also see Freeman (2003; 2007) and Hill (2014).

Paul Weithman (2010) argues that the reason why Rawls turned from *A Theory of justice* to *Political Liberalism* is that Rawls realized that his original argument for the stability of justice as fairness in *A Theory of Justice* failed. And so, according to Weithman, the primary task of Rawls's *Political Liberalism* is to explain how his political conception of justice is stable for the right reason given the fact of reasonable pluralism.

Hyunseop Kim (2015) argues for the stability interpretation of Rawls's law of peoples. He argues that the stability interpretation is better than its alternatives. This interpretation, he argues, explains why Rawls does not opt for the global distributive principles of justice. For since, according to Rawls, domestic political culture is conducive to stability, the global distributive principles of justice are not needed. Kim further argues that this interpretation explains why Rawls insists on the duty of assistance and the toleration of decent peoples. He argues that the assistance from liberal and decent peoples helps burdened societies to become well-ordered, so that they will not be aggressive and undermine the stability of the law of peoples. For if decent peoples are not tolerated by liberal peoples, their self-respect will be hurt and this will undermine the stability of the law of peoples.

Now my Rousseauvian interpretation of Rawls emphasizes the importance of stability in Rawls's early and later works. First, I argue that Rawls's justice as fairness meets the demand of amour-propre, and that once people recognize that their demand of amour-propre is met by justice as fairness, they will willingly comply with the principles of justice. This explains why Rawls's well-ordered society of justice as fairness is stable for the right reason. Second, I argue that Rawls's law of peoples meets the demand of amour-propre, and that once peoples recognize that their demand of amour-propre is met by the law of peoples, they will willingly comply with the principles of the law of peoples. This explains why Rawls's law of peoples is stable for the

right reason. Thus, this difficulty faced by the Kantian interpretation can also be resolved from a Rousseauvian perspective.

To sum up, I have argued that Rawls's Kantian interpretation faces two difficulties both of which can be responded to from a Rousseauvian perspective. The first difficulty is that the Kantian interpretation fails to explain why acting from the two principles of justice is acting autonomously. The second difficulty is that it fails to emphasize the importance of stability. I have argued that this first difficulty is resolved if we understand Rawls's Kantian interpretation as being grounded in Rousseau's (and so also Kant's) idea of autonomy. I have also argued that the second difficulty can be resolved because Rousseau's concept of amour-propre (correctly interpreted by Kant) plays a significant role in Rawlsian stability.

3. The Rousseauvian Interpretation of Rawls

In this dissertation, I argue that there is a Rousseauvian interpretation of Rawls's political philosophy. As pointed out by many commentators, there are a number of important ways in which Rawls's works are indebted to Rousseau, for example, the idea of natural goodness of humans, the idea of the social contract, the idea of the general will, the idea of a realistic utopia, the idea of the legislator and so on.[19] But I will focus on just one way in which Rawls's work is indebted to Rousseau, one way that has not received adequate attention, namely, Rousseau's

[19] For example, Samuel Freeman (2007) points out that Rousseau's idea of the natural goodness of humans, his idea of equal rights of political participation, and his idea of the general will have important influences on Rawls. Frederick Neuhouser (2014) argues that Rousseau and Rawls share some similarities concerning how social cooperation is to be arranged to achieve justice. He argues that Rawls's idea of self-respect as the most important primary good can be traced back to Rousseau's idea of amour-propre. He also argues that Rawls's fair equality of opportunity and fair value of equal political rights both converge with Rousseau's aim of the absence of domination through equality in the political domain. Jeffrey Bercuson (2014) argues that Rawls's ideas parallel Rousseau's in four aspects. First, they both think that institutions can change human nature. Second, they both agree that self-interest can enhance the stability of a society. Third, they both take men as they are and laws as they can be. And finally, they both seek the reconciliation between self-interest and the common good.

moral psychology, especially his concept of amour-propre. I argue that my Rousseauvian interpretation of Rawls merits special attention because it explains away some of the difficulties facing other Rousseauvian interpretation of Rawls.

3.1 Other Rousssseauvian Interpretations of Rawls and Their Shortcomings

There are many themes that enter into the Rousseauvian interpretation of Rawls. For example, first, Rawls and Rousseau both think that persons are self-interested and that they are capable of the sense of justice. Second, they both recognize that institutions can change human nature. Third, Rawls's idea of public reason parallels Rousseau's idea of the general will. Fourth, similar to Rousseau, Rawls's idea of realistic utopia takes men as they are and laws as they can be. Fifth, they both have the idea of a legislator who performs a decisive role in the transformation of our social nature. I will begin by introducing the first three themes of the Rousseauvian interpretation and the difficulties they face. In the next section, I will introduce my Rousseauvian interpretation, and argue that it can resolve these difficulties.

First, as Samuel Freeman (2007) points out, Rawls largely shares the same picture of human nature with Rousseau. They both reject the Christian idea of original sin and the Hobbesian account of human nature which characterizes human beings as purely self-interested. They both think that human beings are capable of compassion and the sense of justice. And they both think that political and social institutions have great impact on the kind of person one is. But one might argue that it is not clear why a person who is self-interested is capable of the sense of justice. For example, G. A Cohen (2008, p.178) argues that Rawls has an incoherent account of human nature, and that Rawls's rational agents are unjustly selfish. Freeman's interpretation does not give us a response to Cohen's objection.

Second, it is worth emphasizing that Rawls and Rousseau both think that social and political institutions can change human nature. Jeffrey Bercuson (2014) focuses on this theme in his Rousseauvian interpretation of Rawls. He argues that Rousseau's claim that "man is naturally good, and it is through institutions alone that men become bad" (p.68) should be read in a Rawlsian (and also Hegelian) way: first, "human nature is reconcilable with a legitimate and stable (i.e., egalitarian) system of social and political institutions;" second, "institutions exercise an important — no, decisive — influence on the character of human beings: individuals take their cue from the political institutions that coerce them, and so we ought not expect a people to be anything other than what their institutions make them" (ibid). Based on this reading, he argues that for both Rousseau and Rawls, since we have the fundamental interests of freedom and self-respect, and since these interests can be satisfied by the egalitarian social and political institutions, our self-interest is reconcilable with political stability and the common good. But Bercuson's interpretation mainly focuses on the common aspects of the Hegelian and the Rousseauvian interpretations of Rawls, that is, the institutional impacts on human nature, and so the significance of Rawls's concept of amour-propre is underestimated.

Third, Fredrick Neuhouser (2008) points out that Rousseau's idea of the general will parallels Rawls's idea of public reason. For Rousseau, people have their own judgments about public matters, but to see if their judgments are correct, they must appeal to the general will. If their judgments conform with the general will, their judgments are correct; if their judgments do not conform with the general will, their judgments are incorrect. Neuhouser argues that this suggests that Rousseau has the idea of public reason: it is not his own reason, but public reason that tells him what he ought to do in the public sphere. Neuhouser thus concludes that Rousseau's idea of the general will parallels Rawls's idea of pubic reason. But one might argue that Neuhouser does

not provide an adequate explanation for why Rousseau's idea of the general will can be associated with Rawls's idea of public reason. For example, David James (2011) argues that since pity plays a significant role in the formation of the general will — it inclines people to understand the interests of others, Rousseau's idea of the general will has no parallel with Rawls's idea of public reason, in the formation of which pity plays no role.

In the next section, I will introduce my Rousseauvian interpretation of Rawls and argue that these difficulties others face can be resolved by this interpretation.

3.2 My Rousseauvian Interpretation of Rawls

My Rousseauvian interpretation focuses on the significance of amour-propre in Rawls's early and later works. I argue that Rousseau's concept of amour-propre has a positive form, namely, the desire to have equal standing with others. This reading of amour-propre, what Rawls calls "the wide view of amour-propre," is different from the traditional reading of amour-propre. The traditional view argues that amour-propre only has a negative form, namely, the desire to be superior to others. I argue that the traditional reading does not treat Rousseau's works as a consistent whole. The traditional reading of amour-propre cannot explain the consistency between Rousseau's second *Discourse* and the *Social Contract*. It is also not supported by Rousseau's moral psychology in *Emile*. Based on this reading, I argue that Rousseau's concept of amour-propre can serve as a starting point for the correct understanding of Rawls. Similar to Rousseau, Rawls aims to describe a well-ordered society that publicly affirms the equal standings of each of its citizens. Rawls agrees with Rousseau that once everyone's equal standing is publicly recognized by the basic structure of the society, they are provided with the social bases of self-respect (amour-propre). Now the public affirmation of the two principles of

justice provides the social bases of self-respect (amour-propre) in Rawls's justice as fairness. Once citizens in the well-ordered society know that their social bases of self-respect are ensured by the two principles of justice, they will willingly comply with the two principles of justice. This will bring about the stability of the well-ordered society of justice as fairness.

In *A Theory of Justice*, Rawls argues that if citizens in the well-ordered society of justice as fairness have acquired the sense of justice in Stage 3 of moral development, and if the sense of justice is in congruence with the thin conception of the good, and if such a society does not generate special psychologies such as envy, the well-ordered society of justice as fairness is stable for the right reason. Now, since Rawls's moral psychology is an illustration of Rousseau's thesis that "the sense of justice is no mere moral conception formed by the understanding alone, but a true sentiment of the heart enlightened by reason, the natural outcome of our primitive affections" (CP 96), and since Rawls's congruence argument in *A Theory of justice* and his overlapping consensus argument in *Political liberalism* are both grounded in Rousseau's conception of persons, and since Rawls's argument for the absence of special psychology is grounded in Rousseau's concept of amour-propre, Rawls's stability argument parallels, and is grounded in, Rousseau's thoughts on amour-propre.

Rawls's stability argument is not complete unless he has shown that the well-ordered societies can sustain their stability in the face of external threats. Similar to the domestic case, Rawls argues, in *The law of Peoples*, that the social bases of self-respect (amour-propre) of both liberal and decent well-ordered peoples are ensured by the public affirmation of the equal standing of each member of the law of peoples. Since their equal statuses are recognized by the law of peoples, they will willingly comply with the principles of the law of peoples. This will bring about the stability of the law of peoples.

I argue that my Rousseauvian interpretation of Rawls can resolve some of the difficulties facing other Rousseauvian interpretations. Firstly, it resolves the difficulty Freeman's Rousseauvian interpretation faces, that is, how a person who is self-interested is capable of the sense of justice. According to my interpretation, amour-propre has both a positive and a negative form. Once people have the positive form of amour-propre, they are ready to grant others' equal standing with them. This reciprocity of disposition is manifested in every stage of moral development, and it explains why we will acquire the sense of justice once we know that we and those around us are beneficiaries of the just institution, provided we have gone through the first two stages of moral development. Thus, Rousseau's concept of amour-propre explains why a person who is self-interested is capable of the sense of justice.

Secondly, my interpretation goes beyond the common aspects of the Hegelian and the Rousseauvian interpretations of Rawls that Bercuson focuses on. My interpretation emphasizes the significance of amour-propre in Rawls's early and later works. It explains how amour-propre makes a just society possible, and how a just society transforms the nature of amour-propre. I argue that amour-propre can be easily inflamed once we enter society. Once our amour-propre is inflamed, we will be in a relation of domination and subjection, a relation of exploitation and subjugation with others. The only way to cure these ills is to have a political institution that recognizes each and all as free and equal beings. One of the common aims of Rousseau and Rawls is to describe a society of mutual respect and mutual regard. In the well-ordered society of justice as fairness, since the social bases of self-respect are ensured by the public affirmation of the two principles of justice, if every citizen has gone through the three stages of moral development, they will have a secure sense of self-respect that inclines them to respect others. Thus, on the one hand, Rousseau's and Rawls's political institutions explain the transformation

of our social nature, namely, amour-propre; on the other hand, the transformation of amour-propre explains how a society of mutual respect and mutual regard is possible.

Thirdly, my interpretation explains why Rousseau's idea of the general will is associated with Rawls's idea of public reason. I argue that amour-propre is the primary motivating force behind Rousseau's just society. And so the reason why individuals in the state of nature willingly subject themselves to the general will is that the demand of amour-propre and the demand of amour de soi are met by Rousseau's society. Amour-propre is also the primary motivating force for Rawls's well-ordered society of justice as fairness. And so the reason why citizens in the well-ordered society of justice as fairness willingly comply with public reason is that the demand of amour-propre and the demand of amour de soi are met by Rawls's justice as fairness. Thus, my interpretation explains why Rousseau's idea of the general will parallels Rawls's idea of public reason.

To sum up, I have argued that the Freeman's Rousseauvian interpretation of Rawls does not provide an explanation for how, in both Rousseau and Rawls, a person who is self-interested is capable of the sense of justice. Moreover, the Bercuson's Rousseauvian interpretation of Rawls does not emphasize the significance of amour-propre in Rawls's early and later works. Finally, the Neuhouser's Rousseauvian interpretation does not provide an adequate explanation for why Rousseau's concept of amour-propre parallels Rawls's idea of public reason. I have argued that these difficulties can be resolved by my Rousseauvian interpretation. I have argued that the first difficulty is explained away by both Rousseau's and Rawls's moral psychology. The second difficulty is resolved by my interpretation which focuses on the significance of amour-propre in Rawls's early and later works. The third difficulty is explained away if amour-propre plays a significance role in both the formation of the general will and that of public reason.

4. The Task of Each Chapter, and How It Contributes to the Overall Project

Chapter 1: I aim to depict the general picture of my Rousseauvian interpretation of Rawls and compare it with other interpretations. I argue that my interpretation of Rawls can resolve some of the difficulties faced by other interpretations. First, I argue that the Kantian interpretation of Rawls faces two difficulties that can be resolved by my interpretation. The first difficulty is that it fails to explain why, for Rawls, acting from the two principles of justice is acting autonomously. The second difficulty is that it fails to emphasize the importance of stability in Rawls. I argue that the first difficulty is resolved because Rawls's justice as fairness is grounded in Rousseau's (and so also Kant's) idea of autonomy, and that the second difficulty is resolved because Rousseau's concept of amour-propre (correctly interpreted by Kant) plays a significant role in Rawlsian stability. Second, I argue that other themes that factor into the Rousseauvian interpretation face several difficulties that can be resolved by my interpretation. First, it explains away the difficulty facing Freeman's interpretation, that is, how a person who is self-interested is capable of the sense of justice. Second, it explains away the difficulty facing Bercuson's interpretation, that is, how significant the role of amour-propre is in Rawls. Third, it explains away the difficulty facing Neuhouser's interpretation, that is, why Rousseau's idea of the general will parallels Rawls's idea of public reason.

Chapter 2: I aim to defend Rawls's interpretation of Rousseau's concept of amour-propre. The traditional view of amour-propre is that amour-propre only has a negative form, namely, the desire to be superior to others. This view, however, contradicts Rousseau's own interpretation. The wide view of amour-propre, endorsed by Dent, Rawls, Cohen and Neuhouser, argues that amour-propre also has a positive form, namely, the desire to have equal standing with others. However, Bloom would argue that the traditional view is consistent with Rousseau's works

because it is not amour-propre but compassion that makes the society of mutual respect and mutual regard possible. I argue that Bloom's objection is implausible because for Rousseau compassion is too inert to motivate human actions. I argue that, for Rousseau, self-love or self-respect (amour de soi or amour-propre) is the primary motivating force, as is evident in the second *Discourse*, *Emile* and other of Rousseau's political works. Furthermore, I argue that the wide view of amour-propre presents us with a Rousseau that is distinct from natural law theorists and utilitarians. This also explains why Rawls shifts from the natural law tradition and the utilitarian tradition to the social contract tradition after reading Rousseau.

Chapter 3: I aim to show that there is a Rousseauvian interpretation of Rawls's justice as fairness by focusing on the significance of Rousseau's concept of amour-propre in Rawls's justice as fairness. I argue that for Rawls, one of the main reasons why parties in the original position would choose justice as fairness over other principles of justice is that justice as fairness meets the demand of amour-propre, namely, the need for having equal status with others, while other principles do not. For Rawls, justice as fairness meets the demand of amour-propre because the public affirmation of equal citizenship provides the social bases of self-respect. He argues that classic and average utilitarianism (or perfectionism) fail to provide the social bases of self-respect because they allow the sacrifice of one's freedom for the greater good of many (or the maximization of human excellences). In explaining why Rawls's justice as fairness meets the demand of amour-propre, I also argue that the assumptions of Rawls's argument from self-respect for the two principles of justice parallel, and are grounded in, Rousseau's thoughts on amour-propre. I argue that Rawls's assumption that self-respect is perhaps the most important primary good parallels Rousseau's story of an unhappy fugitive who has no self-esteem. His assumption that the public affirmation of equal citizenship provides the social bases of self-

respect that makes a society of mutual respect and mutual regard possible also parallels, and is grounded in, Rousseau's thoughts on amour-propre, especially the relative nature, the conditions and the reciprocal feature of amour-propre. I conclude that there is a Rousseauvian interpretation of Rawls's justice as fairness.

Chapter 4: I aim to show that there is a Rousseauvian interpretation of Rawls's stability argument. There are three parts of Rawls's stability argument: his moral psychology, the congruence argument in *A Theory of Justice* or the overlapping consensus argument in *Political Liberalism*, and his argument from the absence of special psychologies. First, I try to explain why Rawls's moral psychology is an illustration of Rousseau's thesis. I argue that the reason for this is that, for Rawls, the natural feelings of love and friendship are necessary for the acquisition of the sense of justice. I also argue that Rawls's moral psychology parallels, and is grounded in, the reciprocal feature of amour-propre. I also argue that Rawls's claim that his moral psychology is more realistic than other conceptions of justice parallels, and is grounded in, Rousseau's claim that amour-propre is the primary motivating force to stabilize the well-ordered society. Second, I argue that Rawls's congruence argument in *A Theory of justice* is grounded in the Rousseauvian and Kantian conception of moral autonomy. I also argue that Rawls's overlapping consensus argument in *Political Liberalism* is grounded in Rousseau's conception of political autonomy and his idea of public reason. Third, I argue that Rawls's argument from the absence of special psychologies also parallels Rousseau's concept of amour-propre because meeting the demand of amour-propre is the main reason why Rawls's well-ordered society of justice as fairness does not generate special psychologies such as envy.

Chapter 5: I aim to show that there is a Rousseauvian interpretation of Rawls's law of peoples by focusing on the significance of amour-propre in Rawls's law of peoples. First, I argue

that Rousseau's concept of amour-propre plays a significant role in three themes of Rawls's law of peoples, that is, the two international original positions, the idea of a realistic utopia, and the distinction between peoples and states. Second, I argue that amour-propre plays a significant role in Rawls's argument that his law of peoples meets the demand of amour-propre, his stability argument in the law of peoples, and his arguments against other principles of international justice such as political realism and distributive cosmopolitanism. I then consider two objections against Rawls's argument that his law of peoples meets the demand of amour-propre. The first objection is that Rawls's law of peoples neglects women's need for recognition in decent peoples; the second objection is that Rawls's law of peoples does not grant equal status to burdened societies, outlaw states and benevolent absolutism. I argue that there are Rousseauvian responses to these objections. In response to the first objection, I argue that Rawls's treatment of decent peoples meets the demand of amour-propre for all peoples and it meets the demand of amour-propre at least in a minimal sense for all natural persons. In response to the second objection, I argue that Rawls's treatment of outlaw states and other societies meets the demand of amour-propre for all natural persons, and that outlaw states and other societies are not moral persons that have a positive sense of amour-propre. I conclude that there is a Rousseauvian interpretation of Rawls's law of peoples. In particular, Rousseau's concept of amour-propre plays a significant role in Rawls's law of peoples.

Chapter 6: I have assumed that the legitimate social bases of self-respect are equal basic rights and liberties and their substantive fair value. Given this assumption, I have argued that Rawls's justice as fairness and his law of peoples meet the demand of amour-propre. I consider two objections against this assumption. The first objection comes from some Chinese nationalists who argue that Chinese people's self-respect comes from their achievements and

accomplishments in their history, especially the traditional culture of China. The second objection is from luck egalitarians who argue that a person's self-respect comes from economic equality. I argue that these two objections are groundless.

CHAPTER 2 ROUSSEAU'S CONCEPT OF AMOUR-PROPRE

Rousseau is one of the most enigmatic writers in the history of philosophy. Commentators have different opinions about whether he is an individualist or a collectivist, a rationalist or an irrationalist, a socialist or a non-socialist, a pessimist or an optimist, a Deist, a Catholic or a Protestant, etc. Since, in each of these debates, both sides can marshal arguments from Rousseau's writings to support their view, some commentators have concluded that Rousseau is an inconsistent and incoherent writer. But this verdict underestimates the integrity of Rousseau's works. It also fails to interpret Rousseau as he interprets himself: "All that is daring in the *Contrat Social* had previously appeared in the *Discours sur l'inégalité*; all that is daring in *Emile* had previously appeared in *Julie*" (Cassirer & Gay, 1954, p.3). I take this principle as paramount in interpreting the work of Rousseau, that is: there is nothing incoherent or inconsistent in his works as a whole. On the basis of this principle, I argue that the traditional view of Rousseau's concept of amour-propre fails, and that the wide view advocated by Nicholas Dent (1988; 1998), John Rawls (2007), Joshua Cohen (1997) and Frederick Neuhouser (2008; 2014) succeeds.

Amour-propre is sometimes translated as vanity or pride,[20] sometimes as self-love or self-esteem.[21] Commentators generally agree Rousseau's use of amour-propre is in contrast with amour de soi, which is usually understood as a kind of self-love or self-esteem that is independent of others and which refers to the desire for self-preservation or for one's well-being.

---

[20] Christopher Bertram (2004, p.22) says that amour-propre is "often somewhat misleadingly translated as pride or vanity."

[21] Allan Bloom (1979, p.11) says that amour de soi is a "natural and healthy self-love and self-esteem," and amour-propre is "a self-love relative to other men's opinion of him". But as we will later see, Rousseau explicitly says that amour-propre can also have a natural or healthy form.

But it is a matter of controversy what amour-propre refers to. The traditional view of Rousseau's concept of amour-propre takes it to be exclusively negative, that is, the desire to be superior to others.[22] According to this view, in the state of nature, one's amour de soi can be satisfied by meeting one's basic needs, and so men are self-sufficient and free; but in society, one's amour de soi is replaced by amour-propre, which can never be satisfied, and so men become unfree and unhappy. Rousseau's psychological remedy to this ill, according to this view, is to return to amour de soi by freeing one from the enslavement of amour-propre. But the opponents of this traditional view of amour-propre argue that amour-propre also has a positive aspect.[23] For example, Dent, Rawls, Cohen, and Neuhouser argue that the traditional view fails to understand Rousseau's works as a whole. They point out that this view seems incompatible with Rousseau's views in the *Social Contract* and in *Emile*. If amour-propre manifests solely negatively, what provides the psychological foundation for the society described in the *Social Contract*? Furthermore, this view neglects some of Rousseau's insights in *Emile,* where Rousseau explicitly suggests that amour-propre can have a positive form. Proponents of the traditional view argue that their reading of amour-propre is also compatible with Rousseau's other works, since, in their view, it is compassion that provides the psychological foundation for the society Rousseau describes in the *Social Contract,* not amour-propre. It is this wide view of amour-propre that I defend in this chapter; I argue that amour-propre has a positive aspect, i.e., the desire for public esteem, or the desire to have, or to be evaluated as having equal status with others.[24] I argue that

---

[22] For, one of the most recent proponents of this reading, see Pauline Chazan (1993).
[23] For example, Grace Roosevelt (1941), Roger Masters (1968), Judith Schklar (1969), Nicholas Dent (1988; 1998), Joshua Cohen (1997), Timothy O'Hagan (1997;1999), Christopher Bertram (2004), David Gauthier (2006), John Rawls (2007), Mark Button (2008), Christopher Brooke (2010), Niko Kolodny (2010), Jeffrey Bercuson (2014), and Frederick Neuhouser (2008; 2014).
[24] The desire for public esteem is different from the desire for superiority in the sense that the former can be satisfied by one's having or being evaluated as having equal status with others.

for Rousseau the psychological foundation of a just society is not pity but amour-propre in its healthy form, as is clear from his second *Discourse*, *Emile*, and other political works. I argue that this thesis is a great principle in Rousseau's political works, without which we cannot fully understand Rousseau. On the basis of this principle, we are in a position to read Rousseau as a distinctive figure in the history of political philosophy, one whose theory of a just society is different from that of natural law theorists and utilitarian theorists. Furthermore, presumably, Rawls anticipated the wide view of amour-propre in mid-1950s, which explains his transition from the natural law tradition and the Millian utilitarian tradition to the social contract tradition.

1. The Traditional View of Amour-propre

Rousseau's *Discourse on the Origin of Inequality* is addressed to the question "what is the origin of (moral) inequality?" In the beginning of the second *Discourse*, he draws a distinction between natural inequality, consisting of differences in age, health, strength of the body and quality of the mind, and social inequality (or what he calls moral inequality), consisting of differences in fame, power, wealth, etc. He argues that the first kind of inequality is justified because it will not give rise to relations of domination and subjection, but whether the second kind of inequality can be justified or not requires careful examination. What is the origin of social inequality? Rousseau argues that original human nature is made of original dispositions and capabilities from nature alone. The original dispositions are self-love (amour de soi) and pity; the capabilities from nature are perfectibility and free will. Since social inequality originated neither in the original dispositions nor in the capabilities from nature, he argues that social inequality does not originate in human nature. Therefore, moral inequality must originate

from somewhere other than human nature. According to the traditional view, Rousseau proposes that this inequality originates in an artificial passion that seeks social advantages, or what is called "amour-propre." The traditional view argues that with the appearance of amour-propre, social inequality emerges;[25] with the development of agriculture and metallurgy, social inequality is enlarged; the institutionalization of private property enhances the inequality between the rich and the poor; and with the appearance of political institutions to protect the advantage of the rich, another kind of social inequality emerges, the inequality between the powerful and the powerless. Under this regime, the poor and powerless are forced to sell themselves for their subsistence. The rich and the powerful exploit the less advantaged while they are dominated or alienated by their insatiable desire for prestige, wealth and power. Therefore, everyone in society is in a relationship of domination and subjugation, alienation and subjection, exploitation and suppression. Thus, according to the traditional view, amour-propre is the desire to be superior to others that makes men harm each other and compete with each other. Or as Rousseau writes:

> Finally, consuming ambition, the ardent desire to raise one's relative fortune less out of genuine need than in order to place oneself above others, instills in all men a black inclination to harm one another, a secret jealousy that is all the more dangerous as it often assumes the mask of benevolence in order to strike its blow in greater safety: in a word, competition and rivalry on the one hand, conflict of interests on the other, and always the hidden desire to profit at another's expense;

[25] The first appearance of amour-propre is occasioned by the development of human intelligence, the get-together of idle men and women, and so on. (DI 166)

> all these evils are the first effect of property, and the inseparable train of nascent inequality. (OW, II, 171)

Following the passage quoted above, according to the traditional interpretation, it is natural to interpret amour-propre as exclusively negative, "the ardent desire to raise one's relative fortune less out of genuine need than in order to place oneself above others."

On this interpretation, it is natural to interpret Rousseau as expressing in *Emile* a yearning to return to amour de soi by freeing oneself from the enslavement of amour-propre, as Rousseau says:

> Self-love, which regards only ourselves, is contented when our true needs are satisfied. But *amour-propre*, which makes comparisons, is never content and never could be, because this sentiment, preferring ourselves to others, also demands others to prefer us to themselves, which is impossible. This is how the gentle and affectionate passions are born of self-love, and how the hateful and irascible passions are born of *amour-propre*. (E 213-4)

The educative purpose of *Emile*, according to the traditional view, is to identify Emile's social self with his natural self through negative education or natural education. In other words, according to this reading, by preventing the emergence of amour-propre, Emile can be his natural, free and happy self, even in a corrupted society.[26] The principles guiding the earliest stage of Emile's moral development seem to support this reading. For example, the first principle is to "let [children] have all the strength that nature gives them-—a strength they would not know how to abuse" (E 68); the second principle is to "aid and supplement what is lacking to them,

[26] As we will see later, the educational plan in *Emile* stands in stark contrast to the educational plan of Polish people in *Considerations on the Government of Poland*.

whether in intelligence or strength, in all that is connected with physical need" (E 68); the third is to "limit oneself solely to the really useful, without granting anything to whim or to desire without reason" (E 68); and the fourth is to "distinguish in their desires what comes immediately from nature and what comes from opinion" (E 68). Thus, according to the traditional view of amour-propre, Rousseau's educative purpose in *Emile* is to return to amour de soi. On this reading, if one is merely moved by amour de soi in society, his needs will not exceed his strength and he will be self-sufficient, free, and happy, just like he was in the state of nature.

2. The Wide View of Amour-propre

The traditional view of amour-propre maintains that Rousseau holds the thesis that amour-propre is exclusively negative, i.e., the desire to be superior to others, and that it follows from this thesis that Rousseau's educative purpose in *Emile* is to return to amour de soi. But if this reading is right, Rousseau is an incoherent and inconsistent writer, for he says in *Rousseau juge de Jean-Jacques* that:

> human nature does not go backward, and it is never possible to return to the times of innocence and equality once they have been left behind. This too is one of the principles on which he has most insisted. So that his object could not be to bring populous peoples or great States back to their first simplicity, but only to stop, if it were possible, the progress of those whose small size and situation have preserved from such a swift advance toward the perfection of society and the deterioration of the species. These distinctions deserved to be made and were not. He was stubbornly accused of wanting to destroy the sciences, the Arts, the theaters, the Academies and to plunge the universe back into its first barbarism; and on the

> contrary he always insisted on the preservation of existing institutions, holding that their destruction would only move the palliatives while leaving the vices and substituting brigandage for corruption. (RJJ 213)

Here Rousseau says that he never advocated human beings to go back to the barbarous state of nature; rather, he always insisted that the cure to the ills in the existing institutions lies in the institutions themselves. Thus if we accept Rousseau's interpretation of himself, we must reexamine the traditional view of amour-propre, which claims that amour-propre cannot but be inflamed when men step into society. The opponents of the traditional view of amour-propre provide a new reading of amour-propre which tries to render Rousseau's works as a coherent and consistent whole.

In Dent's *Rousseau* (1988), he argues that Rousseau's idea of amour-propre has great force in explaining human psychology and shaping today's world, a world in which domination and subordination, mastery and subjection still widely exist. Rousseau's aim, he argues, is to fight "the nightmarish version of man and society" (Dent, 1988, p.5), and to do that, citizens need to respect each other's equal status and have full participation in the legislative activities of a society. But what is the psychological foundation of a society of mutual respect and equality? If one's major drive in society is merely a desire for distinction, it is impossible to realize the ideal of mutual respect and equality. For this reason, but not only this reason, Dent argues there must be a positive form of amour-propre in human psychology that makes the ideal of mutual respect and equality possible.

Dent admits that Rousseau focuses on the negative form of amour-propre in his early writings. In the *Discourse on the Sciences and Arts*, Rousseau criticizes people's high achievements in artistic and scientific fields as being mostly motivated by their desires for

distinction and for prestige, not by their passions for truth. In the *Discourse on the Origin of Inequality*, he writes that the origin of inequalities is amour-propre, a desire for recognition and acknowledgement for oneself from others, an artificial feeling that arises in peoples' hearts when they step into society. But amour-propre must also have a positive form that provides the psychological foundation for the society Rousseau describes in the *Social Contract*. This form of amour-propre is most explicit in Book IV of *Emile*, where Rousseau says:

> Since my Emile has until now looked only at himself, the first glance he casts on his fellows leads him to compare himself with them. And the first sentiment aroused in him by this comparison is the desire to be in the first position. This is the point where love of self turns into amour-propre and where begin to arise all the passions which depend on this one. But to decide whether among these passions the dominant ones in his character will be humane and gentle or cruel and malignant, envy and covetousness, we must know what position he will feel he has among men, and what kinds of obstacles he may believe he has to overcome to reach the position he wants to occupy. (E 235)

From this, as Dent argues, it follows that the demand of amour-propre can be met if we know what position we have among men, that is, a position of mutual regard and respect, a position of equality (cf. Dent 1988, p.105). In other words, our desire for having, or being evaluated as having equal status with others can be met if we know that we have equal status with others. Thus, since I do not need to dominate you to satisfy my amour-propre, the satisfaction of my amour-propre does not have to be in conflict with yours; rather, the satisfaction of my amour-propre requires the satisfaction of yours. Thus granting that you have equal status with me is the requirement of my desire for having equal status with you. In this

sense, self-respect and respect for others are mutually supportive. In other words, if one wishes to acquire human recognition and respect for oneself, one must acknowledge others' status as equal with oneself (ibid., p.149). Dent argues that Rousseau's *Social Contract* describes a society in which everyone is recognized as having equal status with everyone else, that is, equal citizenship. Thus, the equality granted to everyone in such a society satisfies a person's amour-propre which, in turn, provides for the possibility of a society of mutual regard and mutual respect.

Rawls accepts Dent's interpretation of amour-propre for two reasons. First, he argues that Kant also interprets Rousseau's amour-propre as having a positive form, as Kant says:

> The predisposition to humanity can be brought under the general title of self-love which is physical and yet compares… that is to say, we judge ourselves happy or unhappy only by making comparisons with others. Out of this self-love springs the inclination to acquire worth in the opinion of others. This is originally a desire merely for equality, to allow no one superiority above oneself, bound up with a constant care lest others strive to obtain such superiority; but from this arises gradually the unjustifiable craving to win it for oneself over others.[27]

Here Kant interprets Rousseau's amour-propre as "the inclination to acquire worth in the opinion of others" and "a desire merely for equality" (ibid). Rawls says that it is only through Kant's interpretation of amour-propre that he finally understands Rousseau's second *Discourse*. Similar to Dent's view, the second reason Rawls gives is that the wide view of amour-propre is

---

[27] Immanuel Kant, *Religion within the Limits of Reason Alone*, trans. Theodore M. Greene and Hoyt H. Hudson (La Shalle, II: Open Court, 1934), bk.1, sec. 1. This passage is quoted by Rawls (LHPP 199) in support of the wide view of amour-propre. Dent (1992, pp.35-6; 1998, pp.105-6; 1998, pp.65) and Cohen (1997, pp.134-35) also connect this passage with Rousseau's concept of amour-propre.

necessary for understanding Rousseau's works as coherent and consistent. If amour-propre is merely an artificial passion to be superior to others, what is the psychological foundation for the political society depicted in the *Social Contract*? If there is no psychological force that motivates people to achieve that kind of society, the political society depicted in the *Social Contract* is utterly utopian and Rousseau's works are generally pessimistic. Therefore, Rawls says that if we agree that Rousseau is a coherent and consistent writer, we should accept that amour-propre has a positive form.

Cohen agrees that if amour-propre is only a desire to be superior to others, then it is impossible for Rousseau to address the motivation problem, namely, to specify what motivates one to conform to the general will. He argues that one solution to the motivation problem lies in the natural form of amour-propre, i.e., "a concern that they affirm my worth" (Cohen, 1997, p.109). This form of amour-propre is natural because when I develop a relation with others as I enter society, self-love, which includes a sense of self-worth, naturally extends to include a concern for others' affirmation of my self-worth. Amour-propre in its natural form, therefore, Cohen argues, can be understood as a concern that others value me as I value myself. This interpretation of amour-propre, Cohen says, is supported by Rousseau, who says: "This is, then, the summary of the whole human wisdom in the use of the passions: (1) To have a sense of the true relations of man, with respect to the species as well as the individual. (2) To order all the affections of the soul according to these relations" (E 219). Here, Rousseau explicitly says that in contrast to the false relations of man — competition and rivalry — there are "the true relations of man" with respect to "the species as well as the individual." What is the true relation of man? Rousseau later points out that "Man is the same in all stations. If that is so, the stations having the most members merit the most respect. To the man who thinks, all the civil distinctions

disappear. He sees the same passions, the same sentiments in the hodcarrier and the illustrious man" (E 222). Amour-propre interpreted as having a natural form, therefore, captures the true relations of man, i.e., the equal stations of man. Properly understood, amour-propre should have a natural or positive form, that is, the desire for public esteem, or the desire to have or to be evaluated as having equal status with others.

Neuhouser (2008; 2014) also argues that amour-propre has both a negative form and a positive form. On one hand Neuhouser agrees with Dent, Rawls and Cohen that if amour-propre only has a negative form, it is impossible for the various evils diagnosed in the *Second Discourse* to be remedied; one the other hand, he provides a different interpretation for Note XV in the second *Discourse* from commentators who take Rousseau to be suggesting that amour-propre is exclusively negative.[28] In Note XV, Rousseau says:

> Amour propre [vanity] and Amour de soi-même [self-love], two very different passions in their nature and their effects, should not be confused. Self-love is a natural sentiment which inclines every animal to attend to its self-preservation and which, guided in man by reason and modified by pity, produces humanity and virtue. Amour propre is only a relative sentiment, factitious, and born in society, which inclines every individual to set greater store by himself than by anyone else, inspires men with all the evils they do one another, and is the genuine source of honor. (DI, 218)

Commentators usually take this distinction between amour de soi and amour-propre as evidence in supporting of the traditional interpretation of amour-propre.[29] But, as Neuhouser

---

[28] See, for example, Norma Thompson (2000, p.74); Pleshette DeArmitt (2013, pp.31-2); and Pablo Muchnik (2009, pp.144-5).
[29] Ibid.

points out, Rousseau does not mean to say here that the inclination "to set greater store by himself than by anyone else" is the necessary consequence of amour-propre, but rather that it is the likely consequence; Rousseau does not mean to say that amour-propre is the source of all the evils regardless of its forms, but that when amour-propre is occasioned with certain circumstances, such as leisure, luxury, division of labor, etc., it is the source of all the evils.[30] It follows from Neuhouser's reading that the traditional reading of amour-propre misunderstands Rousseau. The meaning of "the ardent desire to raise one's relative fortune less out of genuine need than in order to place oneself above others, instills in all men a black inclination to harm one another" is that when amour-propre is inflamed it instills in all men an inclination to harm others (OW, II, 171). Furthermore, when Rousseau hyperbolically says "*amour-propre*, which makes comparisons, is never content and never could be" (E 213-4), he means that amour-propre is likely to be insatiable, but will not necessarily be so.

3. An Argument Against the Wide View of Amour-propre

As discussed above, the opponents of the traditional view of amour-propre, for example, Dent, Rawls, Cohen and Neuhouser, argue that amour-propre has a positive form because the traditional view is incompatible with Rousseau's works as a whole. But the proponents of the traditional view of amour-propre might further argue that the traditional view of amour-propre is well supported by Rousseau's *Emile* and that it is compatible with Rousseau's interpretation of himself.[31] For example, in Allan Bloom's introduction to *Emile*, amour-propre, a "keystone of Rousseau's psychological teaching" (Bloom, 1979, p.10), is said to be the source of all ills, such

[30] Christopher Bertram (2004, p.22-3) agrees that Rousseau's Note XV in the second *Discourse* does not support the traditional reading of amour-propre, even though it appears to.
[31] See, for example, Chazan (1993, p.348) and Bloom (1979).

as "anger, pride, vanity, resentment, revenge, jealousy, indignation, competition, slavishness, humility, capriciousness, rebelliousness" (ibid., p.11). Bloom argues that Rousseau's intention of negative education here is to "prevent amour de soi from turning into amour-propre, for this is the true source of man's dividedness" (ibid., p.10), and that government and law are needed to prevent the emergence of amour-propre. Therefore, if Bloom's interpretation is plausible, amour-propre is not really the psychological foundation of society described in the *Social Contract,* but rathe the source of inequality and human ills.

What is the psychological foundation of the society described in the *Social Contract*, if not amour-propre? Is Rousseau's project in the *Discourse on the Origin of Inequality* compatible with his project in the *Social Contract*, if amour-propre does not have a positive form? These problems, Bloom seems to suggest, can be resolved if we take Rousseau's ideas of compassion into account.[32] Unlike amour-propre, according to Bloom, compassion is a natural feeling in human hearts. Bloom suggests that Rousseau's view is that compassion, rather than amour-propre, makes a society of mutual respect and mutual regard possible. He writes:

> The recognition of our sameness and our common vulnerability dampens the harsh competitiveness and egotism of egalitarian political orders. Rousseau takes advantage of the tendency to compassion resulting from equality, and uses it, rather than self-interest, as the glue binding men together. (ibid., p.19)

The purpose of Emile's education, Bloom argues, is to let Emile realize the miseries of human kind and nurture his feelings of compassion.[33] The first level of Emile's education in

---

[32] Jonathan Marks (2007) agrees with Bloom that Rousseau takes compassion as the basis of a just society.
[33] It is interesting that Rousseau merely talks about the nurture of compassion only in Book IV of *Emile*. If as Bloom suggests, the educative purpose of *Emile* is to nurture his compassion to free Emile from the enslavement of amour-propre, why doesn't Rousseau discuss the role of compassion from the start? Rousseau fully acknowledges the

compassion is to let Emile study arts instead of science. This study will enable Emile to despise the vanity of Plutarch's heroes and pity their unfortunate ends. The second level is to study Robinson Crusoe's island and to produce "contempt for the great of this world, not a slave's contempt founded in envy, indignation, and resentment, but the contempt stemming from a conviction of superiority which admits of honest fellow feeling and is the precondition of compassion" (ibid., 19). The third level of education is to study fables and satires, which make Emile identify himself with the sufferers and have a feeling of compassion toward human beings. If the education is successful, he suggests, Emile's compassion, together with his desire for self-preservation, becomes his principle of action, according to which, "concern for others becomes part of his sense of his own interest" (ibid., 20).

If Bloom's interpretation, that is, that compassion or pity, instead of the positive form of amour-propre, is the psychological foundation of the political society depicted in the *Social Contract*, is plausible, perhaps the traditional reading of amour-propre is not mistaken in saying that amour-propre is exclusively negative. But as I will argue presently, in agreement with Dent, Rawls, Cohen and Neuhouser, the traditional view of amour-propre fails to understand Rousseau's works as a coherent and consistent whole. This implies that compassion cannot be the psychological foundation of the political society of mutual respect and equality. In fact, if the ground of morality, for Rousseau, is the natural feeling of compassion, his political society of mutual respect and equality is merely a utopia and his political theory is pessimistic. As Rawls points out, this reading makes Rousseau's overall thoughts unworkable and thus fails to be persuasive.

---

weakness of compassion. He describes the state of nature as a state of peace, not because of the natural pity of human beings, but because of their amour de soi, i.e., their needs that does not exceed their strength.

4. My Defense of the Wide View of Amour-propre

Bloom suggests that compassion can serve as the psychological foundation of the society described in the *Social Contract*. Rousseau, however, never says that compassion is the basis of a just society.[34] Rather, for Rousseau, amour de soi is the primary motivating force in the state of nature. And amour-propre is the primary motivating force in the civil state. In other words, self-love or self-esteem is the primary motivating force in the state of nature and in the civil state. This is not to say that compassion plays no role in the state of nature or in the civil state. In the state of nature, compassion is one of the principles of action and one of the reasons why human beings are naturally good. In the civil state, compassion plays an important role in preventing amour-propre from being inflamed. But as we shall see, compassion is not the primary motivating force for people in the state of nature or in the civil state.

4.1 The Role of Pity vs the Role of Self-love or Self-esteem in the Second *Discourse*

In the *Discourse on the Origin of Inequality*, Rousseau defines compassion as "a natural repugnance to seeing any sentient Being, and especially any being like ourselves, perish or suffer" (DI 127). The feeling of compassion and the desire for self-preservation (amour de soi) are the two principles of action in the state of nature and the foundation of the rules of natural right. But compassion has less motivational force than amour de soi (self-love or self-esteem). As Rousseau says:

[34] Shklar (1969, p.47), Viroli (1988, pp.49-50), and Neuhouser (2014, p.39) agree that, for Rousseau, compassion is too weak to be the basis for a just society.

> as long as he does not resist the inner impulsion of commiseration, he will never harm another man or even any sentient being, except in the legitimate case when, his preservation being involved, he is obliged to give himself preference. (DI 127)

Here Rousseau explicitly says that when pity is in harmony with one's desire for self-preservation, it will motivate man not to harm other animals, especially other human beings; but when one's feeling of pity is in conflict with one's desire of self-preservation, he will give priority to his desire for self-preservation. Later in the same *Discourse*, he criticizes Hobbes's natural philosophy and argues:

> By reasoning on the basis of the principles he establishes, this Author [Hobbes] should have said that, since the state of Nature is the state in which the care for our own preservation is least prejudicial to the self-preservation of others, it follows that this state was the most conducive to Peace and the best suited to Mankind. (DI 151)

Here Rousseau makes clear that he agrees with Hobbes that the primary principle of action in the state of nature is amour de soi (self-love or self-esteem), especially, the desire for self-preservation. But he disagrees with Hobbes that this principle necessarily leads to conflicts of interests and a state of war between one and all; rather, he argues that in the state of nature one's concern for self-preservation is rarely in conflict with the self-preservation of others. This is the reason why the state of nature is a state of peace. For Rousseau, men in the state of nature do not know what is good and are ignorant of vices. And so it is impossible that men in the state of nature are, as Hobbes contends, naturally wicked. Once one's desire for self-preservation is satisfied, he does not care whether others have more fortunes than he and so his desire for self-preservation is rarely in conflict with the self-preservation of others. But as one steps into

society, he develops a sense of self-love or self-esteem that is dependent on the opinion of others, that is, amour-propre. This relative sense of self-love or self-esteem can be inflamed under certain circumstances. Motivated by inflamed amour-propre, everyone seeks social advantages and superiority, and thus the civil state is a state of war. In the state of war, compassion has no motivational force against (inflamed) amour-propre. Rousseau explicitly says that compassion is "developed but weak in Civil man" (DI 153). This weakness of compassion is vividly presented by Rousseau: "One of his kind can with impunity be murdered beneath his window; he only has to put his hands over his ears and to argue with himself a little in order to prevent Nature, which rebels within him, from letting him identify with the man being assassinated" (DI 153). He later argues that this inflamed amour-propre which stifles natural pity "made men greedy, ambitious, and wicked" (DI 171). Therefore, self-love or self-esteem, rather than pity, as Rousseau suggests in the second *Discourse*, is the primary motivating force both in the state of nature and in the civil state.

4.2 The Role of Pity vs the Role of Self-love or Self-esteem in *Emile*

The importance of self-love or self-esteem is also explicit in *Emile* where Rousseau argues that the purpose of the earlier stages of moral education is to satisfy the child's self-love or self-esteem (amour de soi) and to free him from the enslavement of inflamed amour-propre. In Book I, as we have discussed, the education for the infant stage of moral development is to aid or supplement one's strength to meet his natural needs and to limit his desires within his strength. Thus, the primary motivating force in this stage is self-love or self-esteem (amour de soi), that is, one's desire for self-preservation. In Book II, the educative purpose in the child stage of moral development is to free him from the enslavement of the negative form of amour-propre, and to

nurture the love of freedom in the child's heart. Rousseau says of the fundamental maxim in this stage that "The truly freeman wants only what he can do and does what he pleases" (E 84). In Book III, the purpose of the adolescent stage of moral development is to teach useful knowledge, such as geometry, geography, astronomy, physics, social studies and carpentry, which contributes to one's well-being which consists in "health, freedom, and the necessities of life constitute it" (E 177). Thus in a corrupted society, like the one in which Emile lives, before he enters society, the purpose of education is to satisfy his self-love or self-esteem (amour de soi), which consists in desire for self-preservation and concern for one's own well-being, and freeing one from the enslavement of the inflamed amour-propre. The role of pity is negligible in the first three stages of moral development.

In Book IV of *Emile,* Rousseau describes Emile's adulthood. As a person enters society, he is in a relation with other human beings. His amour de soi develops into amour-propre. Amour-propre can be good and useful, as Rousseau says:

> The sole passion natural to man is amour de soi or amour-propre taken in an extended sense. This amour-propre in itself or relative to us is good and useful; and since it has no necessary relation to others, it is in this respect naturally neutral. It becomes good or bad only by the application made of it and the relations given to it. Therefore, up to the time when the guide of amour-propre, which is *reason*, can be born, it is important for a child to do nothing because he is seen or heard-nothing, in a word, in relation to others; he must respond only to what nature asks of him, and then he will do nothing but good. (E 93)[35]

[35] Italics added.

Here Rousseau explicitly says that amour-propre in its natural form is good and useful. He summarizes the principle of education in the earlier stages of moral development as negative education, that is, "to do nothing." He also suggests that the purpose of education in the adult stage of moral development is to guide amour-propre with reason. The role of reason can vary for Rousseau. In the second *Discourse*, he says: "We seek to know only because we desire to enjoy, and it is possible to conceive why someone who had neither desires nor fears would take the trouble to reason" (DI 142). But he also suggests that: "Self-love is a natural sentiment which inclines every animal to attend to its self-preservation and which, guided in man by *reason* and modified by pity, produces humanity and virtue" (DI 218).[36] Some commentators interpret Rousseau as a sentimentalist who wants to overturn "the Platonic order that gives primacy to reason over feeling" (Kelly, 1997, p.30; Qvortrup, 2003, p.115). But this interpretation ignores the force that Rousseau attributes to reason. Properly understood, there are two kinds of reason for Rousseau, means-ends reason and ends-in-itself reason. The first kind of reason is conducive to the realization of a person's desires. The second kind of reason can guide a person's animal desires. The second kind of reason also explains why for Rousseau persons are autonomous beings, as he says: "for the impulsion of mere appetite is slavery, and obedience to the law one has prescribed for oneself is freedom" (SC 54).[37] Thus, one of the central tasks of Emile's education is to enable him to be an autonomous person whose reason guides his amour-propre into its healthy and natural form. It is worth exploring how reason guides amour-propre into a natural form. Niko Kolodny (2010) provides an explanation:

---

[36] Italics added.

[37] As Neuhouser (2011, p.483) points out: "autonomy is free choice carried out in accordance with a conception of one's own essential identity, a self-conception that supplies the autonomous being with reasons for endorsing some of its desires as worthy of being acted on and for rejecting others as incompatible with one's essential nature."

> If amour-propre is inflamed by the illusion that one's value depends on one's social advantage, then healthy amour-propre would seem to lie in *independent* knowledge of one's true value, as consisting in something *other than social advantage*. (Kolodny, 2010, p.192)

Kolodny correctly points out that a necessary condition for the natural or healthy form of amour-propre is that one knows his true value independent of his social advantage. But this is not an adequate explanation for the natural or healthy form of amour-propre, which Kolodny defines as "A desire to have, and to be evaluated by all others as having a certain value in comparison with all others, including at least moral equality" (ibid., 170). Even though one knows his true value is independent of his social advantage, if he believes others have the illusion that his value depends on his social advantage, he will still desire to have, and to be evaluated by all others as having, social advantage. This point follows from Rousseau's definition of (natural or inflamed) amour-propre as a relative sentiment that depends on the opinion of others. Therefore, the necessary conditions for the healthy or natural form of amour-propre is (1) the knowledge that one's true value is independent of one's social advantage; and (2) the knowledge that others know that one's true value is independent of one's social advantage. But these two conditions do not suffice to explain the healthy or natural form of amour-propre because they merely explain why one's amour-propre is not inflamed, not why one's amour-propre is natural or healthy. To know where the natural or healthy amour-propre lies is to know what is good and useful, as Rousseau says: "This amour-propre in itself or relative to us is good and useful" (E 93). The fundamental good for human beings is freedom, as Rousseau points out:

> The only one who does his own will is he who, in order to do it, has no need to put another's arms at the end of his own; from which it follows that the first of all

> goods is not authority but freedom. The truly freeman wants only what he can do and does what he pleases. That is my fundamental maxim. (E 84)

Thus it is clear that Rousseau thinks that the natural or healthy form of amour-propre lies in (1) the knowledge that one's true value is independent of one's social advantage; and (2) the knowledge that others know that one's true value is independent of one's social advantage; and (3) the knowledge that one's true value, i.e., the utmost value of all values, is freedom (and equality);[38] and (4) the knowledge that others know that one's true value is freedom (and equality). Therefore, properly understood, the natural or healthy form of amour-propre is the desire for freedom and equality.

The above argument constitutes an important part of my main argument that pity cannot provide the psychological foundation for the political society depicted in the *Social Contract*. For it establishes the possibility for the natural form of amour-propre as the psychological foundation for the political society of mutual respect and equality. In the country in which Emile lives in, if one is properly educated like Emile, he will have a sort of amour-propre in its natural form because he knows one's true value is freedom (and equality), not social advantage.[39] It is thus Emile's duty to live among them and "cultivate their friendship in sweet association, be their benefactor and their model" (E 474) so that his associates know that one's true value is freedom (and equality). Once people in the country, a small association, know, and know that other associates know, the true value is freedom (and equality), they will have a sort of amour-propre in its natural form. But in a corrupted society, like that in which Emile lives, it is impossible for Emile and his countrymen acquire a robust sense of amour-propre. In a corrupted society, since it

---

[38] For Rousseau, equality is necessary for freedom.
[39] This country is the place where Emile can pursue a patriarchal and rustic life.

is publically known that some people are inferior to others (or simply it is not publically known that everyone is free and equal), people will necessarily have an inflamed form of amour-propre. Those who are advantaged will attempt to satisfy their self-esteem by subjecting others. And those who are disadvantaged will sell their freedom to others for social advantages. Emile and his countrymen will not have an inflamed form of amour-propre for they know that their true value lies in freedom and equality. But they will not have a robust sense of amour-propre in its natural and healthy form either for they know that it is not public knowledge that a person's true value lies in freedom and equality in the corrupted society to which that they belong.

4.3 The Role of Pity vs the Role of Self-love or Self-esteem in Rousseau's Other Works

The educational goal of Emile is to enable him to guide his amour-propre with reason so that he knows that his true value lies in freedom and equality. But this task is a difficult one in the corrupted society in which Emile lives. Even though the tutor's education makes him able to acquire the knowledge that his true value lies in freedom and equality, considering that his self-esteem would not be supported by the society that he belongs to, it is likely that Emile would only have a less than robust sense of self-esteem (amour-propre). In the *Political Economy*, Rousseau provides another way of guiding amour-propre with reason, as he says: "his own reason should be suspect to him, and he should follow no other rule than *public reason*, which is the law" (PE 5).[40] Here the reason that guides amour-propre is not individual reason as we have discussed above, but public reason or intersubjective reason. In the *Social Contract*, Rousseau

[40] Italics added.

seems to suggest that this public reason is the general will: by examining whether a person's own reason conforms with the general will, he knows whether his own reasoning is correct or false.[41] Rousseau says: "Each one states his opinion on this by voting, and the declaration of the general will is drawn from the counting of the votes. Therefore when the opinion contrary to mine prevails, that proves only that I was mistaken, and what I thought to be the general will was not" (SC 124). Now we understand why Rousseau says that the fundamental problem that the social contract tries to solve is "To find a form of association that will defend and protect the person and goods of each associate with the full common force, and by means of which each, uniting with all, nevertheless obey only himself and remain as free and before" (SC 49-50). By submitting himself to the general will, a person submits his own reason to the public reason. And since the general will (or the public reason) is a part of his own will (or reason), he submits to nothing but himself by submitting to the general will (or the public reason). Since freedom is the obedience to the law one gives oneself, a person is as free as before. Furthermore, since the social contract is an agreement among equals, everyone knows and knows that others know that everyone totally alienates himself to the general will. In other words, everyone knows, and knows that others know, that his true value lies in his freedom and equality which he acquires through the social contract. Therefore, in the *Social Contract*, the way for amour-propre to manifest its natural and healthy form is to guide it with public reason, or the general will.

But the starting point of the social unity of the general will is not compassion. What motivates human beings to form the association at issue is their concern for self-preservation and well-being, and their concern for freedom and equality. Unless a person's own amour-propre is satisfied, he is in no position to pity others. The extension of a person's amour-propre starts with

---

[41] For this interpretation of Rousseau's idea of reason, see Neuhouser (2008, pp.201-13).

the people around him. This is also why Diderot's cosmopolitan ideal, "the general will of human race" is merely utopian. In Rousseau's argument against Diderot concerning the origin of the general will, he claims that the concepts of universal or international justice must first be nurtured in the context of a small political community (cf. Roosevelt, 1941, p.125). Diderot regards "the general will of human race" as the ground of peace and justice (cf. ibid., p.133), but for Rousseau this idea is utterly utopian, as Rousseau says:

> Every particular society, when it is narrow and unified, is estranged from the all-encompassing society. Every patriot is harsh to foreigners. They are only men. They are nothing in his eyes. This is a drawback, inevitable but not compelling. The essential thing is to be good to the people with whom one lives. (E 39)

The precondition of the love of humankind is the love of country which originates from the love of oneself. Compassion, therefore, is the end of peace and justice, not the starting point.[42] Rousseau does not trust "those cosmopolitans who go to great length in their books to discover duties they do not deign to fulfill around them" (E 39). The purpose of public education in a small community, e.g., Poland in the eighteen century, is, for Rousseau, to nurture a patriotic zeal in the hearts' of its people, as he says in the *Considerations on the Government of Poland*: "Every true republican has drunk in love of country, that is to say love of law and liberty, along with his mother's milk. This love is his whole existence" (Roosevelt, 1941, p.134). A nation's amour de soi is its desire for self-preservation. When this amour de soi develops into an inflamed amour-propre, as Rousseau writes in his "Summary" of the Abbé de Saint-Pierre's *Project for Perpetual Peace*, an empire has a tendency to encroach upon its neighbors (cf. ibid., p.131). But

[42] This point is made by Cassirer (1954). He says: "A form of sympathy which transcends mere egotism may be the goal of society, but it cannot be its point of departure" (p.102).

this amour de soi can develop into a natural or healthy amour-propre as in the case of Poland, as the love of country, the love of law and liberty (cf. ibid., p.132).[43] The natural development of a small community thus is analogous to the natural development of a person. In the early stages of its development, it is only moved by amour de soi, especially, its concern for self-preservation; gradually, as it develops relations with its neighbor countries, the amour de soi matures into amour-propre, which can manifest in both a negative and a positive form. Under the guidance of reason, that is, through the public education which nurtures the love of country in its citizens, amour-propre manifests in a natural or healthy form, the desire for freedom and equality, or the desire for justice.

5. Conclusion and Some Afterthoughts

To conclude, amour-propre has both a positive form and a negative form. This reading of amour-propre allows us to view Rousseau's works as a coherent and consistent whole. As we have seen in the second *Discourse*, *Emile*, and his other political works, amour-propre in its natural form, that is, the desire for freedom and equality, is the psychological foundation for the political society depicted in the *Social Contract.*

Now we are in a position to read Rousseau as a distinctive figure in the history of political philosophy. First, Rousseau's political theory is different from natural law theorists such as Aristotle, Grotius and Hobbes.[44] Natural law theorists think that the knowledge of what is reasonable can be derived from the knowledge of what is natural for human beings (cf. Bix,

---

[43] For this interpretation of amour-propre, see also Roger D. Masters (1968, p.42); Shklar (1969, p.19); and David Gauthier (2006), chapter 3.

[44] Charles E. Vaughan (1962, pp.16-7), A. Cobban (1934, p. 115, fn), Leo Strauss (1965, p.276) and L.G. Crocker (1968, p.91) correctly points out that Rousseau is never a natural law theorist and that he tries to substitute the concept of natural law with the concept of general will (cf. Gabriella Sivestrini, 2011, p.280).

2004, p.63). Aristotle thinks that natural justice is superior to conventional justice because natural justice is valid for all communities while conventional justice is subject to change (cf. Thomson, 1953, p.138). Grotius and Hobbes also admit the existence of natural law. Rousseau, however, argues that Aristotle's theory of justice justifies slavery; for Aristotle claims that there are natural slaves. Rousseau also argues that for Grotius and Hobbes, "it is an open question whether humankind belongs to a hundred men, or whether those hundred men belong to human kind" (SC 42). He argues that "His [Grotius's] most frequent mode of argument is always to establish right by fact" (SC 42). For Rousseau, it is true that human beings have natural differences in age, health, strength of the body and quality of the mind but these natural differences cannot justify slavery. Rousseau does not think that justice is grounded in nature, for he does not believe that freedom can be obtained by returning to nature. The traditional view mistakenly believes that Rousseau's educational purpose for Emile is to identify his social self with his natural self. But as we have seen, the real purpose of Emile's education is to make him an autonomous person who guides his amour-propre with reason. The domestic education of Emile is similar to the public education of the people in Poland. Only through the transformation of amour-propre, with the guidance of public reason, is it possible to form an association in which people are as free as before. Thus, Rousseau does not belong to the natural law tradition of Aristotle, Grotius and Hobbes. Rather, he argues that the natural law tradition justifies slavery and that it fails to treat human beings as equals.

Secondly, Rousseau's political theory is different from that of Hume and other utilitarian theorists. Hume famously argues that the natural law theorists mistakenly infers "ought" from "is", that is, what is just from what is natural. For Hume, justice is an artificial virtue based on convention. He argues that once people recognize that the political institution to which they

belong is aimed at the public good, they will willingly act in accordance with the rules and the laws set up by the institution. The reason for this is that human beings are capable of taking up the view of the judicious spectator, identifying and sympathizing with the people who are benefited by a political institution (LHPP 186). Therefore, for Hume, it is sympathy that makes a just society possible. Rousseau disagrees with Hume that sympathy is the basis of a just society. As we have seen, for Rousseau, it is self-love or self-esteem, rather than sympathy, that makes a just society possible. In other words, unless we have a secured sense of amour-propre, pity's role is negligible in motivating us to form or retain a just society. Rousseau also disagrees with Hume that the recognition that the political institution is aimed at the public good is sufficient for a sense of justice. Rousseau argues that the virtue of justice originates from self-love (amour-propre) under the guidance of public reason. Therefore, a necessary condition for one's voluntary compliance with rules and laws set up by a political institution is that these rules and laws ensure one's fundamental good, that is, freedom (and equality).

Utilitarianism, like natural law theories, is a teleological theory. Utilitarianism tries to derive what is right from what is good for human beings but in maximizing the good, utilitarianism fails to respect all human beings as equals. Utilitarians, like Hume, argue that sympathy is the psychological basis of a just society. But for Rousseau, even though what is right must take human goodness into considerations, it is never just to subject a person to a lesser liberty for the sake of the greater satisfaction of many. As we have seen, for Rousseau, it is necessary for an association to respect every associate as equal to ensure his or her self-esteem (amour-propre). Utilitarianism thus fails to meet the demand of amour-propre. Furthermore, utilitarianism mistakenly believes that sympathy is the psychological basis of a just society, but as we have

seen, for Rousseau the natural and healthy form of amour-propre is the psychological basis of a just society, not sympathy.

To conclude, the wide reading of Rousseau's concept of amour-propre presents us with an interpretation of a Rousseau whose political theory is different from that of natural law theorists and utilitarian theorists. Rawls might have anticipated the wide view of amour-propre in mid-1950s, and so he says that reading Rousseau with this understanding helps his transition from the natural law tradition and the Millian utilitarian tradition to the social contract tradition (cf. Reidy, 2009, p.36). Since commentators often misunderstand both Rawls and Rousseau, it is important for us to know Rousseau's influence on Rawls so that we can have a better understanding of both philosophers. This is the purpose of my next chapter.

CHAPTER 3 THE SIGNIFICANCE OF ROUSSEAU'S CONCEPT OF AMOUR-PROPRE IN RAWLS'S JUSTICE AS FAIRNESS

In *A Theory of Justice*, Rawls says "there is a Kantian interpretation of the conception of justice from which this principle derives. This interpretation is based upon Kant's notion of autonomy" (TJ 221). Thus, some commentators naturally take Rawls's justice as fairness as mainly influenced by Kant. For example, Stephen Darwall (1980) reads Rawls's *Theory* as a Kantian response to emotivism and utilitarianism in the context of the debate between Humeanism and Kantianism in the last century. But Rawls himself acknowledges that his reading of Rousseau plays a major role in his transition from the natural law and Millian utilitarian tradition to the social contract tradition in the mid-1950s (cf. Reidy, 2009, p.36). And while Rawls criticizes Hobbes, Locke, Kant, Hume, Mill and Marx in his *Lectures*, he makes no critique of Rousseau.[45] In his *Lectures*, as we have seen, Rawls provides an interpretation of amour-propre which is different from the traditional reading. There he argues that amour-propre should be understood as "a natural concern for a secure standing in relation to others and involves a need for equal acceptance with them" (LHPP 198). He also suggests that the positive form of amour-propre makes Rousseau's society, depicted in the *Social Contract,* possible. He later introduces Rousseau's social contract theory, according to which, "we must arrange our political and social institutions according to the terms of cooperation expressed by the social contract" so that our moral freedom, political and social equality and independence can be secured (LHPP 208). Rawls's justice as fairness is also a contract theory, according to which, all parties, free and equal, rational and reasonable, in the

[45] This is pointed out by Robert Jubb (2011, p.247).

Original Position, under the veil of ignorance, would choose the two principles of justice over other conceptions of justice. In the opening pages of his *The Law of Peoples* and *Restatement*, he says that his liberal conception of justice as fairness, which follows Rousseau's thought in The *Social Contract*, describes a realistic utopia which "takes people as they are (by the laws of nature), and constitutional and civil laws as they might be, that is, as they would be in a reasonably just and well-ordered democratic society" (LP 13; JF 13). One of the central tasks of Rawls's justice as fairness, therefore, is to explain how his liberal conception of justice as fairness meets the demand of amour-propre, that is, a person's need for equal status with others, and how this amour-propre in turn makes his well-ordered society in justice as fairness realistic.

In this chapter, I argue that Rawls's justice as fairness has a Rousseauvian interpretation for two reasons. The first reason is that amour-propre (self-respect) [46] plays a significant role in Rawls's justification of the two principles of justice. The second reason is that Rawls's argument from self-respect for the two principles of justice parallels, and is grounded in Rousseau's thoughts on amour-propre.

In the first section, I will briefly introduce Rawls's justice as fairness; in the second section, I will introduce his argument from self-respect for the two principles of justice. As I explain the significance of amour-propre in this argument, I will explain how his assumptions in the argument from self-respect parallel, and are grounded in, Rousseau's thoughts on

---

[46] Amour-propre is properly translated as self-love or self-esteem. Rawls uses self-esteem and self-respect interchangeably in *Theory*. Some commentators such as Stephen Darwall (1977), Nir Eyal (2005, p.201), Robert Yanal (1987, pp.369-70), Jean-Philippe Deranty (2009, p.402), Larry Thomas (1978, p.261-3), argue that Rawls fails to distinguish self-respect and self-esteem while they are distinct in important aspects. Rawls says that self-respect and self-esteem are "not the same" in "A Kantian Conception of Equality (1975)" (CP 260). But he still thinks that self-respect and self-esteem both depend on the equal status and the acceptance of others (cf. James Zink, 2011, p.341).

amour-propre. I conclude that there is a Rousseauvian interpretation of Rawls's justice as fairness.

1. A Brief Introduction to Justice as Fairness

In the opening pages of *A Theory of Justice*, Rawls says that his contract theory of justice aims "to present a conception of justice which generalizes and carries to a higher level of abstraction the familiar theory of the social contract as found, say, in Locke, Rousseau, and Kant" (TJ 10). His social contract theory, like all other social contract theories, has three parts: a description of the state of nature, the contract itself and an outline of the political institutions. He argues that parties who are free and equal rational beings, in the Original Position, under the veil of ignorance, will choose the two principles of justice over other conceptions of justice, such as utilitarianism, perfectionism, and intuitionism. The parties in the original position are seen as free because first, they are capable of having the moral power to have a conception of the good, second, they are self-authenticating sources of valid claims, and thirdly, they are capable of taking responsibility for their ends (JF 21-4; PL 72). The parties are seen as equal because they equally have the two moral powers to a minimal degree to engage in social cooperation (PL 50). They are rational in the sense that they have their own conceptions of the good and they are willing to make efforts and use resources to achieve their plans of life (PL 50). He also presupposes that there are social primary goods that are necessary for the exercise and the development of the two moral powers and the pursuit of any conception of the good. These primary goods are rights and liberties, powers and opportunities, incomes and wealth, and the social bases of self-respect. The Original Position is an initial situation in which free and equal rational beings, who do not know their identities and idiosyncrasies under the veil of ignorance,

try to decide how to distribute these social primary goods. Rawls argues that the parties in the original position would choose the two principles of justice over other conceptions of justice. The parties in the Original Position would not choose intuitionists' principles of justice because intuitionism does not have a priority rule to resolve the conflicts among different first principles of justice. Since the parties are rational, they wish to establish standards according to which their different claims can be weighted. Thus, they would not choose the intuitionists' principles of justice which involve a plurality of conflicting first principles of justice, without a method for weighing these principles against each other. They would not choose the perfectionists' principle of justice because perfectionism takes one conception of the good, that is, human excellence, as exclusively intrinsically valuable and thus cannot be acceptable by all parties in the Original Position.[47] For the parties in the Original Position conceive of themselves and of each other as free to choose and alter their own ends, and so they know that they might not share the same conception of the good with perfectionism. But the perfectionists' society might require them to sacrifice their own freedom for the sake of maximizing perfection. For example, Aristotelian perfectionism notoriously justifies slavery. Thus, the parties in the Original Position would not take the risk of losing their freedom to choose the principle of perfection. For similar reasons, the parties in the Original Position would not choose the utilitarian principle of justice. For utilitarianism might require a person to sacrifice his freedom for the sake of (total or average) utility.[48]

---

[47] The moderate form of perfectionism that Rawls discusses does not take human excellence to be exclusively intrinsically valuable but it faces the same difficulty that strict perfectionism does. For both forms of perfectionism take human excellence to be prior to freedom. (TJ 285-6)

[48] Rawls does not think that restricted utilitarianism faces the same difficulty as utilitarianism does. For restricted utilitarianism thinks that liberties and rights should not be sacrificed for the sake of greater utility (cf. Rawls, 1955). I will discuss these two forms of utilitarianism and their problems later in this chapter. For objections against restricted utilitarianism, see, for example, J. J Smart (1956) and H. J McCloskey (1957).

Rawls argues that the two principles of justice would be chosen by the parties in the Original Position. The two principles of justice are as follows: the first principle of justice, or the principle of equal liberty, says that "Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all" (JF 42); The second principle of justice, which consists of the principle of fair equality of opportunity and the difference principle, says that "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" (JF 42-3). The principles of justice might conflict with each other. If there is no priority rule, there is no way to resolve conflicts. Rawls thus provides several notions of priority to solve this problem: the priority of equal liberty (the first priority rule) which says the first principle of justice is prior to the second principle of justice, the priority of justice over efficiency and welfare (the second priority rule) which includes the priority of the first half (the principle of fair equality of opportunity) of the second principle over the second half (the difference principle) of the second principle, and the priority of second principle over the principle of efficiency and that of maximizing the sum of advantages, and the priority of right over the good.

Why would the parties in the Original Position choose the two principles of justice over other conceptions of justice? One of the main grounds of Rawls's justification for the two principles of justice is the argument from self-respect, which explains how the two principles of justice meet the demand of amour-propre.

2. The Argument from Self-respect for the Two Principles of Justice

In §29 Rawls appeals to the conditions of publicity and finality to ground the two principles of justice. The conditions of publicity and finality are two of the formal constraints on the concept of right. Other formal constraints on the concept of right are the conditions of generality, universality and ordering. These formal constraints on the concept of right must be met if a conception of justice is to be allowed to be presented to the parties in the Original Position. The condition of generality says that principles should be general. And so the egoistic principle that everyone is to serve *my* purpose should not be considered as a candidate for the principles of justice presented to the parties. The condition of universality says that principles should be universal in application. A principle of justice must hold for all moral persons. It cannot be self-contradictory or self-defeating if everyone acts on it. The condition of publicity says that principles should be publicly known to all parties. This condition is met by all contract theories because the word "contract" means that it is accepted by all parties. The condition of ordering says that a conception of justice must be able to adjudicate conflicting claims. The intuitionists' principles of justice do not have a priority rule and thus do not meet the condition of ordering.[49] The condition of finality means that the original agreement is final and there is no second chance to reconsider it. The conditions of publicity and finality are some of the main grounds of the two principles of justice. These two conditions give rise to the strains of commitment, according to which, those who are represented by the parties must honor the principles of justice agreed by the parties in the Original Position. Rawls argues that other conceptions of the good, especially the principle of utility fail to meet the strains of commitment.

[49] They meet the condition of ordering formally, but not substantively. Perhaps it is for this reason it is presented to the parties in the original position to choose from.

On the one hand, if the principle of utility is chosen in the Original Position, parties are to stick to it and act upon it. But for the sake of greater utility, the utilitarian society may require a person to sacrifice his own freedom. Thus if the principle of utility is chosen, by the strains of commitment, a person may need to sacrifice his freedom for the greater good. Since the decision made in the Original Position is final, there is no second chance for them to reconsider the original agreement. And since the agreement is publicly known, persons in the utilitarian society are not motivated to retain the utilitarian society out of the fear that they may be asked to sacrifice their freedom for the sake of the greater satisfaction of many.

I argue that for Rawls what underlies the argument from self-respect for the two principles of justice is the demand of amour-propre. First, I will argue that for Rawls the utilitarian principle fails to meet the demand of amour-propre, the need for equal status, because the utilitarian society undermines self-respect.[50] This is not to deny that utilitarianism affirms one's equal status in an important sense — each counts equally (for one) in the aggregation of utility. But to be publicly affirmed as an equal input to social utility is not the same as being publicly affirmed as a separate equal person within a system of social cooperation. The former undermines one's self-respect while the latter does not. Next, I will argue that for Rawls the two principles of justice meet the demand of amour-propre because the public affirmation of equal citizenship ensures the social bases of one's self-respect. In the mean-time, I will explore some similarities between Rawls's thoughts on self-respect and Rousseau's thoughts on amour-propre. I argue that the premises of Rawls's argument from self-respect parallel, and are grounded in,

[50] Robert Jubb correctly points out that "the discussion of how utilitarianism would fail to meet the strains of commitment is the obverse of the set of demands that the natural form of amour-propre makes" (Jubb, 2011, p.251).

certain of Rousseau's thoughts on amour-propre. If my argument is plausible, there is a Rousseauvian interpretation of Rawls's justice as fairness.

2.1 The Utilitarian Principle Fails to Meet the Demand of Amour-propre

The utilitarian principle that Rawls criticizes for the most part is the principle of utility endorsed by classic utilitaranism such as Bentham, Sidgwick and Edgeworth, and the principle of average utility endorsed by average utilitarianism such as Mill and Wicksell.[51] Classic utilitarianism says that a society is to be arranged so as to maximize the utility. Average utilitarianism says that a society is to be arranged so as to maximize average utility. Rawls argues that both classic utilitarianism and average utilitarianism undermine self-respect. Another kind of utilitarianism he discusses is restricted utilitarianism. Rawls is sympathetic to restricted utilitarianism in his "Two Concepts of Rules (1955)," but in *Theory* and *Restatement* he criticizes mixed utilitarianism, a form of restricted utilitarianism, as indeterminate, and maintains that it fails to meet the condition of publicity (TJ 277-85; JF 126-30). Mixed utilitarianism (that Rawls discusses) says that the society is to be arranged in accordance with the principle of equal liberty and the principle of utility and that the principle of equal liberty is prior to the principle of utility. Rawls suggests that the mixed view is not utilitarian for the principle of utility only has a subordinate place. Since both mixed utilitarianism and justice as fairness agree on the priority of the principle of equal liberty, it is reasonable to postulate that mixed utilitarianism secures self-

[51] Rawls's criticisms against utilitarianism face many criticisms. See, for example, Gregory S. Kavka (1975), Brian Barry (1977), Ivar Labukt (2009), David Lyons (1972), David Braybrooke (1973), Timothy D. Roche (1982), Thomas Carson (1983), Richard Arneson (2000). Most of these criticisms fail to understand Rawls's argument against utilitarianism as the argument from self-respect.

respect to a certain degree.[52] But, as we shall see, both classic utilitarianism and average utilitarianism fail to meet the demand of amour-propre.

2.1.1 Classic Utilitarianism Fails to Meet the Demand of Amour-propre

One of the criticisms Rawls levels against classic utilitarianism is that it fails to "take seriously the distinction of persons" (TJ 24). Classic utilitarianism suggests that "a society is properly arranged when its institutions maximize the net balance of satisfaction" (TJ 21). Classic utilitarianism argues that since a person should sacrifice small pleasure in the present for greater pleasure in the future, a society should sacrifice an individual person's freedom for the sake of the greater good of many. Rawls argues that for this kind of utilitarianism "it does not matter, except indirectly, how this sum of satisfactions is distributed among individuals any more than it matters, except indirectly, how one man distributes his satisfaction over time" (TJ 23). Classic utilitarianism takes the perspective of an impartial spectator who uses "sympathetic identification in guiding our imagination that the principle for one man is applied to society" (TJ 24). This impartial spectator "is the perfect rational individual who identifies with and experiences the desires of others as if these desires were his own" (TJ 24). But classic utilitarianism fails to provide a motivating force for the utilitarian society. Even though from a third person perspective, that is, from the impartial spectator's perspective, every single person is disposable for the sake of the greater good, it is not the case for the persons who live in the utilitarian society. Those who live in the utilitarian society will not be motivated to retain the utilitarian

[52] I will discuss why the priority of equal liberty is crucial for the satisfaction of the demand of amour-propre later in the chapter.

society.[53] Under the condition of finality, if the principle of utility is chosen in the Original Position, they must live upon the principle of utility. Under the condition of publicity, it is publicly known that the utilitarian principle requires a person to sacrifice his own freedom for the greater good. Since everyone in the utilitarian society fears that they might be the one who is disposable, they are not willing to form or retain the utilitarian society. Since it is not rational for citizens in the utilitarian society to retain the society, classic utilitarianism must rely on the role of sympathy or compassion to motivate its citizens to retain the utilitarian society. But "it seems quite incredible that some citizens should be expected, on the basis of political principles, to accept still lower prospects of life for the sake of others" (TJ 154). When a person's self-preservation is at stake, he will not, and legitimately so, give preference to compassion or sympathy. In other words, a person will not, and should not, give up his own freedom for the sake of the greater satisfaction of many. After all, it is *my* freedom that is sacrificed for the greater good of others.

Another important aspect of classic utilitarianism, Rawls argues, is that "In calculating the greatest balance of satisfaction it does not matter, except indirectly, what the desires are for" (TJ 27). But certain conceptions of the good violate our sentiments of justice as they violate our deep conviction in the inviolability of rights and liberties. He writes:

> if men take a certain pleasure in discriminating against one another, in subjecting others to a lesser liberty as a means of enhancing their self-respect, then the satisfaction of these desires must be weighed in our deliberations according to their intensity, or whatever, along with other desires. If society decides to deny

---

[53] Assuming that those who live in the utilitarian society have ordinary human psychology, they will not be motivated to retain the utilitarian society. But we will see that utilitarianism assumes that its people have a peculiar personality.

> them fulfillment, or to suppress them, it is because they tend to be socially destructive and a greater welfare can be achieved in other ways. (TJ 27)

It is important to notice that Rawls suggests here that self-respect has a negative sense. It can be enhanced by "subjecting others to a lesser liberty." This, however, is no surprise to us, for, as we have seen, Rousseau's concept of amour-propre has a negative sense, that is, a desire for superiority, or a desire for domination. Those who have the negative sense of amour-propre thus seek to manipulate or dominate others, to subject others to a lesser liberty. But this negative amour-propre must be suppressed by the society, Rawls argues, regardless of whether its fulfillment will bring about the greater satisfaction of many or not.

What underlies Rawls's objections to classic utilitarianism is the demand of amour-propre to value oneself, or respect for oneself. When a person's freedom is sacrificed for the sake of the greater satisfaction of many, this person is regarded as having inferior status compared with others. If he knows, and knows that others know that he has inferior status compared with others, he will lack a sense of self-respect. Since under the condition of publicity people in the utilitarian society will know, and know that others know, that he has inferior status, they will tend to lack a sense of self-respect. The deferential wife, Uncle Tom, and the self-deprecator are examples of those who lack self-respect (cf. Thomas Hill, 1991, p.5). The deferential wife is a woman who is devoted to serving her husband. Uncle Tom is a Black man who is extremely deferential to his white masters. The self-deprecator is a person who has contempt for himself for he judges that he is not worthy of respect. All of them are examples of servility. These people do not have a secure sense of self-respect. The utilitarian society is similar to the patriarchal society and the white supremacist society in the sense that it encourages the vice of servility. Utilitarianism may justify patriarchy and white supremacy if the husbands and the whites

contribute more to the net balance of satisfaction. In the utilitarian society, the vice of servility might be praised as a virtue to encourage people to sacrifice their freedom for the sake of the greater satisfaction of many. Those who live in the utilitarian society might mistakenly believe that they do not have equal moral status with others. Thus in the utilitarian society the least advantaged people, who mistakenly believe that they have inferior status, will be timorous and lack self-respect, and the more advantaged people, who falsely believe that they have superior status, will be arrogant and self-conceited. Thus the utilitarian society fosters the negative sense of amour-propre, a desire for superiority on the one hand, and the lack of self-respect, a vice of servility on the other. Consumed with the sense of negative amour-propre or a vice of servility, those who live in the utilitarian society can hardly have the moral sensibility that the utilitarian optimistically expects them to have to stabilize the utilitarian society. This society thus is unlikely to be a society of mutual respect and mutual regard, but rather a society of domination and subjection, a society of manipulation and subjugation.

2.1.2 Average Utilitarianism Fails to Meet the Demand of Amour-propre

It is apparent that average utilitarianism faces the same difficulties that classic utilitarianism does for they both takes utility as the satisfaction of desires and they both admit of "interpersonal comparison that can at least be summed at the margin" (TJ 139).[54] If the population of a society remains unchanged, the principle of average utility is equivalent to the principle of total utility. But if the population of a society doubles, the average utility of the society will decrease while the total utility of the society will increase. Rawls argues that parties

[54] Rawls notes that this notion of utility has been largely abandoned by modern economic theory. The new understanding is that utility is a way of representing the choices of economic agents (TJ 143).

in the original position will choose the principle of average utility over the principle of total utility because the parties aim to advance their own self-interest and have no interest in maximizing the sum total of satisfaction (TJ 141).[55]

But the parties in the original position will not choose the principle of average utility over the two principles of justice. Rawls argues that average utilitarianism, just like classic utilitarianism, justifies, "if not slavery or serfdom, at any rate serious infractions of liberty for the sake of greater social benefits" (TJ 135).[56] Since the parties in the original position are assumed to be ignorant of the probability of their chance of turning out to be somebody, they will not take the risk of being subjected to slavery or serfdom by choosing the principle of average utility. Thus, similar to classic utilitarianism, average utilitarianism also fails to meet the demand of amour-propre for it may justify the sacrifice of a person's freedom for the sake of average utility. Rawls acknowledges that average utilitarianism may argue that if the equal liberties are necessary for a secure sense of self-respect, then they should be established because they may promote average utility (TJ 157). But according to the strains of commitment, the principle of average utility should be publicly known by everyone in the utilitarian society and they cannot adopt non-utilitarian principles of justice. Thus, if it happens that the principle of average utility undermines their self-respect, they cannot reconsider the original agreement. Thus, the parties in

---

[55] Some commentators disagree that the parties in the original position will choose the principle of average utility over the principle of total utility. See, for example, Gregory S. Kavka (1975), Brian Barry (1977). Kavka and Barry argues that the parties in the original position know that they exist and this knowledge inclines them to choose the principle of total utility over the principle of average utility because the average utility is higher than the total utility when they do not exist. But there is no reason for Rawls to suppose that the parties in the original position know that they exist or know about the probability of their existence.

[56] David Lyons (1972) argues that utilitarianism does not justify slavery or serfdom more than justice as fairness. But this objection is untenable. As we will see later, even though average utilitarianism may argue that the equal liberties promote average utility, the parties in the original position will still choose the two principles of justice.

the original position will not choose the principle of average utility over the two principles of justice.

Another difficulty that Rawls mentions with average utilitarianism is that it presupposes a bizarre conception of persons. The utilitarian persons do not have any determinate highest-order interest or fundamental ends. They are ready to accept "as defining their good whatever evaluations these rules assign to accept as defining their, or anyone else's, final ends, even if these evaluations conflict with those required by their existing fundamental interests" (TJ 151). Thus, average utilitarianism assumes that those who live in the average utilitarian society will have the moral sensibility to act on and from the principle of average utility even though it may turn out to undermine their own self-respect. Rawls's moral persons, in contrast, have a determinate highest-order interest and fundamental interest. He says that "free persons conceive themselves as beings who can revise and alter their final ends and who give first priority to preserving their liberty in these matters" (TJ 131-2). Thus, Rawls's moral persons, unlike the utilitarian persons, will not accept the principle of average utility which may undermine their fundamental interest. Rawls's moral persons also have moral sensibility, but their moral sensibility, which is based on the principle of reciprocity and equal respect, will not incline them to act on and from the principle of average utility which may undermine their self-respect. As we will see later, the two principles of justice meet the demand of amour-propre, and so Rawls's moral persons are motivated to stabilize the well-ordered society of justice as fairness.

2.2 The Two Principles of Justice Meet the Demand of Amour-propre

As we have seen, parties in the original position, under the veil of ignorance, would not choose the principle of (average) utility because it fails to meet the strains of commitment and

the demand of amour-propre. But why would they choose the two principles of justice? In §29 Rawls argues that "the public recognition of the two principles gives greater support to men's self-respect and this in turn increases the effectiveness of social cooperation" (TJ 154). His argument from self-respect for the two principles of justice therefore goes as follows: (1) self-respect is perhaps the most important primary good; and (2) the public recognition of the two principles gives greater support to a person's self-respect than other conceptions of justice; and therefore (3) the two principles of justice are more defensible than other conceptions of justice. In explaining the reason why the two principles of justice meet the demand of amour-propre, I will argue that the premises of Rawls's argument from self-respect for the two principles of justice parallel, and are grounded in, Rousseau's thoughts on amour-propre. If my argument is plausible, there is a Rousseauvian interpretation of Rawls's justice as fairness.

2.2.1 Self-respect as Perhaps the Most Important Primary Good

First, Rawls argues that self-respect is perhaps the most important primary good. He argues that "A sense of their own worth is necessary if they are to pursue their conception of the good with satisfaction and to take pleasure in its fulfillment," and that "Self-respect is not so much a part of any rational plan of life as the sense that one's plan is worth carrying out" (TJ 154). In §69 He explains what self-respect is and why it is perhaps the most important primary good:

> We may define self-respect (or self-esteem) as having two aspects. First of all, as we noted earlier (§29), it includes a person's sense of his own value, his secure conviction that his conception of his good, his plan of life, is worth carrying out. And second, self-respect implies a confidence in one's ability, so far as it is within

> one's power, to fulfill one's intentions. When we feel that our plans are of little value, we cannot pursue them with pleasure or take delight in their execution. Nor plagued by failure and self-doubt can we continue in our endeavors. It is clear then why self-respect is a primary good. (TJ 386)

Thus, just like other social primary goods, self-respect is an important social primary good because it is desirable for every rational being and it is necessary for the pursuit of any rational plan of life. It is a human good because it is rational for everyone to have self-respect. It is a social primary good because unlike natural primary goods such as health and vigor, intelligence and imagination, it is directly under the control of the basic structure (TJ 54). But why is self-respect perhaps the most important primary good? Or in other words, why is it perhaps more important than rights and liberties, powers and opportunities, income and wealth? Rawls says "Without it [self-respect] nothing may seem worth doing, or if some things have value for us, we lack the will to strive for them. All desire and activity becomes empty and vain, and we sink into apathy and cynicism" (TJ 386). Without other social primary goods, arguably, we may still find our ends worth pursuing, and we may still take pleasure in the pursuit of our own conceptions of the good. But without self-respect, we may not find anything worth doing, or we may not have the confidence to achieve anything. Admittedly, without other social primary goods, especially rights and liberties, it may be extremely difficult for us to achieve any ends of ours. But only if we have self-respect, that is, we believe that our ends are worth pursuing and we believe that we have the ability to achieve them, will we still strive to achieve our plans of life with pleasure. Furthermore, we may say that other primary goods are important primary goods partly because they are social bases of self-respect. Rawls suggests that the equal fair distribution of social primary goods (the two principles of justice) ensures the social bases of

self-respect, and that this is one of the main grounds for the two principles of justice. Thus self-respect is intrinsically valuable while other social primary goods are valuable in part for the sake of self-respect.[57]

Rawls's definition of self-respect bears some similarities to Rousseau's thoughts in Book IV of *Emile*. There Rousseau tells a story of an unhappy fugitive (whom he latter confesses is himself). This young expatriate was born a Calvinist but latter became a fugitive in a foreign land where he changed his religion for bread. He was admitted to an almshouse where he was taught evils and new dogmas. When he wanted to flee, he was locked up. Fortunately, an ecclesiastic came to the almshouse and helped him to escape. Having escaped, the young man soon obtained a fortune. But he lost everything and was left, without a shelter, without food, ready to die. But the ecclesiastic came to him, and tried to help him. He found the young man had no self-respect, and full of self-contempt and contempt for others because he was a victim of opprobrium and contempt. The ecclesiastic helped the young man to gain his self-confidence by respecting him as an equal, and tried to awaken his amour-propre.[58] Rousseau writes:

> He showed him a happier future in the good employment of his talents. He reanimated a generous ardor in his heart by the account of others' noble deeds. In making the boy admire those who had performed them, the priest gave him the desire to perform like deeds. To detach him gradually from his idle and vagrant life, he had the boy make extracts from selected books; and, feigning to need

---

[57] This is not to say that liberties and rights are not intrinsically valuable because they are constitutive of the social bases self-respect. Without liberties and rights, a person can hardly have a secure sense of self-respect, even though he may have an extravagant amount of power or wealth.

[58] Rousseau says that the ecclesiastic treated the young man as an equal: "One day, when he had been given some money to pass out to the poor, the youth, claiming his right as a poor man, was so craven as to ask him for some of it. 'No,' the priest said, 'we are brothers; you are part of me and I ought not to touch this deposit for my use.' Then he gave the youth from his own money as much as he had asked for" (E 264).

> these extracts, he fed the noble sentiment of gratitude in him. He instructed him indirectly by these books. He made the boy gain a good enough opinion of himself so as not to believe he was a being useless for anything good and so as not to want any longer to make himself contemptible in his own eyes. (E 264)

Rousseau's story of the unhappy fugitive tells us that a hostile environment in which the advantaged oppress the disadvantaged fosters a lack of self-respect, shame, and spite. The ecclesiastic's education of the unhappy fugitive is Rousseau's education of Emile in miniature. To rebuild a person's secure sense of self-respect, he must live in an environment in which he is respected, and is known to be respected as equals. He also should have a rational plan of life that is worth carrying out. As he carries out his plan, he will gradually have a stronger sense of self-worth and self-esteem. In the example of the unhappy fugitive, the young man was instructed to make extracts from books. And the ecclesiastic pretended that what the young man did was helpful to him. The affirmation of his self-worth from the ecclesiastic thus enabled him to gain a good opinion of himself, a sense of self-worth, a sense of self-esteem.

2.2.2 The Public Affirmation of Equal Citizenship Ensures the Social Bases of Self-Respect Which Makes a Society of Mutual-Respect Possible

Having established that self-respect is perhaps the most important good, Rawls argues that the public recognition of the two principles gives greater support to a person's self-respect which makes a society of mutual-respect possible. Rawls's argument for this claim has three steps. The first step is the dependency thesis, that is, self-respect depends on the respect of others. In other words, a person's need for equal status can be met only if others grant him equal status. The second step is the legitimacy thesis, that is, the legitimate social bases of self-respect

are the public affirmation of equal citizenship, not equal socio-economic status. In other words, the public affirmation of a person's equal status, equal citizenship, rather than equal socio-economic status, meets his need for equal status. The third step is the reciprocity thesis, that is, self-respect and the respect for others are mutually supportive. Thus self-respect makes a society of mutual respect and mutual regard possible. I will argue that the dependency thesis is grounded in the relative nature of amour-propre, and that the legitimacy thesis parallels the conditions of amour-propre, and that the reciprocity thesis is grounded in the reciprocal feature of amour-propre. Therefore, there is a Rousseauvian interpretation of Rawls's justice as fairness.

2.2.2.1 The Dependency Thesis

To argue for the main thesis that the public affirmation of equal citizenship ensures the social bases of self-respect, Rawls relies on the dependency thesis, that is, self-respect depends on the respect of others. The dependency thesis explains why the problem of the social bases of self-respect is not a private matter, but a public matter, a matter that needs to be, and can be addressed in the public sphere by the government.[59] As he says: "Now our self-respect normally depends upon the respect of others. Unless we feel that our endeavors are respected by them, it is difficult if not impossible for us to maintain the conviction that our ends are worth advancing" (TJ 154). The problem of the social bases of self-respect is a matter that needs to be addressed in the public sphere by the government because it cannot be satisfied by me alone, assuming my self-respect depends on the respect of others. This problem is a matter that can be addressed in the public sphere by the government because the public affirmation of my equal citizenship

[59] Some commentators argue that self-respect is a private matter. See, for example, Robert E. Lane (1982) who quotes Karl Popper in support of this claim.

ensures the social bases of my self-respect, assuming my self-respect depends on the respect of others.

But it may not be clear for some people why self-respect depends on the respect of others. Some commentators argue that Rawls mistakenly assumes self-respect depends on the respect of others.[60] Bird (2008) argues that if we imagine an encounter between a humiliator and the humiliated, the humiliator has the power to undermine the self-respect of the humiliated. But he argues that the victims can remind themselves that their self-respect does not depend on the respect of others and that in this way they can keep their self-respect intact. But Bird fails to see why it is difficult, if not impossible for the victims to deceive themselves that their self-respect does not depend on the respect of others. Rousseau's thoughts on the relative nature of amour-propre explain why it is difficult, if not impossible for them to do so.

Rawls's dependency thesis is grounded on the relative nature of amour-propre. Rousseau suggests that it is necessary and inevitable that our amour-propre depends on how others value us. In his second *Discourse*, Rousseau writes:

> As ideas and sentiments succeed one another, as the mind and the heart grow active, Mankind continues to grow tame, contacts expand and bond ties tighten. It became customary to gather in front of the Huts or around a large Tree: song and dance, true children of love and leisure, became the amusement or rather the occupation of idle men and women gathered together. Everyone began to look at everyone else and to wish to be looked at himself, and pubic esteem acquired a price. The one who sang or danced best; the handsomest, the strongest, the most

[60] For example, Colin Bird (2008) argues that a person's self-respect should not depend on others.

> skillful, or the most eloquent came to be the most highly regarded, and this was the first step at once toward inequality and vice. (DI 166)

Here Rousseau says that with the development of our intelligent faculties and our sentiments, as we have more leisure time and a variety of amusements, we start to have tightened social bonds with each other. And thus our amour-propre becomes relative to others: we can have a sense of self-respect only when we are respected by others. If Rousseau is right that as we associate with each other, our sense of amour-propre relies on what we are in the eyes of others, then it is necessary and inevitable that the respect of oneself depends on the respect of others. If it is necessary and inevitable that self-respect is dependent on the respect of others, the stoic strategy to escape public opinions, as Bird suggests the humiliated should, is but a sort of self-deception. And if self-respect depends on the respect of others, under social institutions where a person is often disrespected by others, he will necessarily lack self-respect. This implies that the cure for the ill lies in the social institutions themselves. A person's need for equal status can only be met when the social bases of self-respect are ensured.

2.2.2.2 The Legitimate Social Bases of Self-Respect

As we have seen, one of the reasons that Bird's argument against Rawls's dependency thesis fails is that he fails to see why the humiliated cannot deceive herself that self-respect does not depend on the respect of others. Another reason his argument fails is that he fails to see that as Rousseau and Rawls both observe the legitimate social bases of self-respect is not power but the public affirmation of equal citizenship. Actually, Bird's objection only confirms the importance of the public affirmation of equal citizenship as the social bases of self-respect. Indeed, if the social base of self-respect is power, the humiliated will have less self-respect due

to lack of power. But if the humiliated knows, and knows that others know, that the social base of self-respect is not power, but equal citizenship, she will have a secure sense of self-respect despite the fact that she has no power. Bird fails to see the institutional factor that underlies a person's secure sense of self-respect. And thus it is difficult for him to understand why the respect of others is owed to the humiliated despite the fact that she is less privileged in a society.

Why can the need of equal status not be met by the public affirmation of socio-economic status such as power status? In §82 Rawls argues that if a society places the social bases of self-respect on the socio-economic status, it will necessarily undermine the self-esteem of its people. Therefore, the legitimate social bases of self-respect are (should be) equal citizenship. Since the priority of equal liberty acknowledges the legitimate social bases of self-respect while other conceptions of justice do not, he concludes that justice as fairness is more defensible than other conceptions of justice.[61] The reason why he thinks so is as follows. Since self-respect depends on the respect of others, if society is to provide the social basis of one's secure sense of self-respect, it must provide an environment in which one's equal worth is recognized and accepted by others. Rawls argues that the well-ordered society provide such an environment: "in a well-ordered society the need for status is met by the public recognition of just institutions, together with the

---

[61] Nir Eyal (2005), Russell Keat and David Miller (1974), Kai Nielsen (1979), Jeanne Zaino (1998) argue that his argument does not provide grounds for the priority of liberty; on the contrary, it lends support for the priority of socioeconomic equality. Robert Taylor (2003) argues that Rawls's self-respect argument can only prove that political liberty has high priority. Robert Lane (1982) and Richard Sinopoli (1994) argue that social or political institutions should not be designed to ensure one's self-respect. Lane argues that self-esteem is a scarce good and should be earned by individuals and that it is not the task of government to maximize the self-esteem of its population. But as we have seen, since self-respect depends on the respect of others, the social bases of self-respect need to be, and can be addressed in the public sphere by the government. Sinopoli argues that Rawls mistakenly assumes that the considerations of equal respect always outweigh the considerations of associational freedom. He argues that in certain circumstances, associations should have the freedom to associate only with a certain kind of person. For example, if a group of lesbians form a company, they should have the freedom not to hire heterosexual men or women. Sinopoli argues that under these circumstances, government should not interfere with their associational freedom. I tend to think that Sinopoli is right to point out that government should not interfere with the associational freedom of these lesbians, but I do not think their associational freedom necessarily contravenes the principle of equal respect.

full and diverse internal life of the many free communities of interests that the equal liberties allow" (TJ 477). Under the condition of publicity, in the well-ordered society everyone accepts and knows that everyone has equal liberties and rights which allow for a large variety of communities in which a person can develop his interests and talents. The equal distribution of fundamental liberties and rights provides a secure standing for one's equal status, and the public affirmation of one's equal status satisfies one's desire for having, and being evaluated as having, equal status with others. Thus Rawls argues that the social bases of self-respect should be the public recognition of one's equal status. This equal status is not equal socio-economic status but the status of equal citizenship because a society which bases one's self-respect on socio-economic status necessarily undermines one's self-esteem, as he says:

> doing this [grounding one's self-respect on socio-economic status] would put them at a disadvantage and weaken their political position. It would also have the effect of publicly establishing their inferiority as defined by the basic structure of society. This subordinate ranking in public life would indeed be humiliating and destructive of self-esteem. (TJ 477)

Here Rawls suggests that if a society takes a person's socio-economic status to be the legitimate source of his self-respect, he will mistakenly believe that his true value lies in social advantages. Thus he argues that the basic structure of the well-ordered society should be arranged in a way that everyone's equal status, equal citizenship is publicly affirmed. If everyone knows, and knows that others know, that their true value lies in something other than social advantages, or more specifically, in one's status of equal citizenship, it is likely that they will

have a healthy and natural form of amour-propre.[62] Notably, Rawls does not neglect the important role of social and economic equality in ensuring the social bases of self-respect. He acknowledges that if social and economic inequality is large enough, it will undermine self-respect and arouse special psychologies such as envy. The equal fair distribution of social primary goods (the two principles of justice) is designed to ensure that social and economic discrepancies between one and another are beneficial to all. But he emphasizes that the principle of equal liberty is prior to the second principle of justice, and that the public affirmation of the priority of equal liberty ensures the social bases of self-respect.

Rawls's thoughts on the legitimate social bases of self-respect parallels the conditions of amour-propre. As we have seen in the previous chapter, Rousseau thinks that the belief that a person's true value lies in social advantages necessarily leads him to have a desire for superiority, a corrupted form of amour-propre. He argues that this corrupted form of amour-propre is the cause of inequalities and other evils in society and that a just society should guide amour-propre with reason so that amour-propre can manifest its healthy and natural form. The reason that guides amour-propre is the public knowledge that one's true value lies in his equal status, equal citizenship. In other words, the necessary condition for a person's healthy and natural amour-propre is the public knowledge that one's true value lies in his equal status, equal

---

[62] Elsewhere, Rawls says that the public affirmation of equal citizenship and its fair value ensures the social bases of self-respect. Meena Krishnamurthy (2013) argues that Rawls has not given any reason why people's equal worth is respected by social institutions only when all individuals have the same political rights and liberties. She tries to develop Rawls's argument from self-respect and explain why equal political rights and liberties are essential for self-respect. She argues that since a secure sense of self-respect is necessary for the exercise and the development of the two moral powers, equal political rights and their fair value must be ensured to secure self-respect. First, she supposes the minority groups in a society do not have rights to vote. This suggests that their moral powers are less valuable than those of others, and thus they will have a less secure sense of self-respect. Thus, these minority groups will not support institutions that do not ensure their equal liberty to vote. Second, she supposes the rich controls the course of legislation for their advantage. This suggests that the poor's determinate conception of the good is not equally valuable as that of the rich, and this will undermine the poor's sense of self-respect.

citizenship. Rousseau's social contract is an agreement among equals. Therefore, it is publicly known and accepted that all citizens have equal status. The social contract among all equals therefore is the remedy for the pervasive desires for domination, as Rousseau writes:

> If there is any means of remedying this ill in society, it is to substitute law for man and to arm the general wills with a real strength superior to the action of every particular will. If the laws of nations could, like those of nature, have an inflexibility that no human force could ever conquer, dependence on men would then become dependence on things again; in the republic all of the advantages of the natural state would be united with those of the civil state, and freedom which keeps man exempt from vices would be joined to morality which raises him to virtue. (E 85)

Rousseau's social contract theory provides a remedy for this ill. By "the total alienation of each associate with all of his rights to the whole community" (SC 50), citizens in the civil state are as free and equal as they were in the state of nature. Since each and all totally alienate themselves to the union, everyone is equal. And thus no one is superior to anyone else. Since this is an agreement by all, it is publicly known that everyone has equal status. Therefore, in a Rousseauvian society, it is unlikely that people will have a negative sense of amour-propre.

2.2.2.3 The Reciprocity Thesis

After Rawls has established that the public affirmation of equal citizenship ensures the social bases of self-respect, he says that since self-respect and the respect for others are mutually supportive (call it the reciprocity thesis), one's secure sense of self-respect makes the society depicted in the *Social Contract*, the society of mutual respect and mutual regard, possible.

Rawls's reciprocity thesis is grounded in the reciprocal feature of amour-propre. In Rawls's *Lecture* on Rousseau, he says that for Rousseau, in the society under the social contract, people are concerned about two things. The first is our own well-being which is the object of our amour de soi. The second is our equal status with others which is the object of our amour-propre. Rawls suggests that for Rousseau our natural form of amour-propre, the desire for equal status with others, makes a rightful claim on our conduct, a claim for us to grant others' equal status with us, as Rawls says:

> on the basis of our needs and wants we can make claims which are endorsed by others as imposing rightful limits on their conduct. Needing and asking for this acceptance from others involves giving the same to them in return. For, moved by this natural amour-propre, we are ready to grant the very same standing to others, and to recognize the rightful limits that their needs and rightful claims impose on us, provided — and this is essential — our equal status is accepted and made secure in social arrangement. (LHPP 198)

According to Rawls in this passage, if our needs for equal status are met by the social arrangement, our desires for equal status require us to grant others equal standing with us. Thus the precondition of a society of mutual respect and equality is the public affirmation of a person's equal status, that is, equal citizenship. This explains why in the well-ordered society citizens have a natural duty for mutual respect or civility. Since everyone rationally desires the respect of others, it is rational for them to grant others equal standing with them which is the precondition of the satisfaction of their own self-respect. This also explains why unreasonable claims must be "overruled" in the well-ordered society. In Rawls's example of the slaveholder who gets great satisfaction from subjecting others to a lesser liberty, the slaveholder does not

have a healthy and natural form of amour-propre, a desire to have equal status with others. Rather, he has a negative sense of amour-propre, a desire for domination, a desire for subjecting others to lesser liberty. But this desire for domination cannot be universalized because as Rousseau says, it is impossible for everyone's negative amour-propre to be satisfied. Thus the desire for domination cannot make rightful claim on our conduct. Rather, as Rawls says, this claim must be "overruled."[63]

3. Conclusion

To conclude, I have argued that amour-propre plays a significant role in Rawls's argument from self-respect for the two principles of justice. The two principles of justice are more defensible than the principle of (average) utility because the latter cannot satisfy the demand of amour-propre. The (classic and average) utilitarian society necessarily undermine one's self-respect because it requires a person to sacrifice his freedom for the sake of (total or average) utility. Rawls's two principles of justice can satisfy the demand of amour-propre because the public affirmation of equal citizenship ensures the social bases of self-respect.

I have also argued that Rawls's argument from self-respect for the two principles of justice parallels, and is grounded, in Rousseau's thoughts on amour-propre. Each premise of Rawls's argument from self-respect parallels, and is grounded in, Rousseau's thoughts on amour-propre: (1) Rawls's self-respect as the most important primary good seems to be inspired by Rousseau's story of the unhappy fugitive, a depiction of Emile's moral education in miniature. (2) There are three reasons why the public affirmation of equal citizenship ensures the social

---

[63] It is not hard to connect Rousseau's thought here with Kant's Formula of Universal Law: "*act only in accordance with that maxim through which you can at the same time will that it become a universal law*" (Kant, 1996, p.73). Adam Cureton (2013) suggests that "Kant proceeds from our own rational self-regard, through our willingness to reciprocate with others, to duties of respect for others" (p.166). If Rawls's interpretation of Rousseau here is right, Kant's theory of respect is likely to have its root in Rousseau.

bases of self-respect which makes the society of mutual respect and mutual regard possible. The first reason, the dependency thesis, is grounded in the relative nature of amour-propre. The second reason, the legitimacy thesis, parallels the conditions of amour-propre. And the third reason, the reciprocity thesis, is grounded in the reciprocal feature of amour-propre. Thus, if my argument is plausible, there is a Rousseauvian interpretation of Rawls's justice as fairness.

On my interpretation, there are several places (§5-6, §26-30, §64, §67, §82) in which Rawls's objections to the utilitarianism and his justifications for the two principles of justice are grounded in Rousseau's thoughts on amour-propre. There are several advantages of my interpretation over other interpretations such as the Kantian interpretation. First, my interpretation explains why Rawls says that after he read Rousseau with understanding, he turned from the natural law tradition and the utilitarian tradition to the social contract tradition. Second, my interpretation explains why Rawls does not criticize Rousseau while he criticizes Hobbes, Locke, Kant, Hume, Mill and Marx in *Lecture*. Third, my interpretation explains why in *The Law of Peoples* and in *Restatement* Rawls says that his liberal conception of justice aims to describe a realistic utopia which takes "men as they are and laws as they might be". Fourth, my interpretation explains why Rawls insists on the wide view of amour-propre and why he suggests that the healthy and natural form of amour-propre makes a society of mutual respect and mutual regard possible. Notably, to say that Rawls's justice as fairness has a Rousseauvian interpretation is not to say that Rawls's justice as fairness is not influenced by other great philosophers such as Kant. My purpose in this chapter is merely to show that there is a Rousseauvian interpretation of Rawls's justice as fairness and thus Rawls's justice as fairness is not distinctively Kantian as Stephan Darwall (1980) and Robert Taylor (2003; 2011) suggest.

CHAPTER 4 THE SIGNIFICANCE OF ROUSSEAU'S CONCEPT OF AMOUR-PROPRE IN RAWLS'S STABILITY ARGUMENT

We have seen that there is a Rousseauvian interpretation of Rawls's justice as fairness for two reasons. First, for Rawls, one of the main reasons why the two principles of justice are more defensible than other conceptions of justice, especially utilitarianism, is that the two principles of justice meet the demand of amour-propre. And second, Rawls's argument from self-respect for the two principles of justice parallels, and is grounded in, Rousseau's thoughts on amour-propre, in particular, its relative nature, its conditions and its reciprocal feature. But Rousseau's influence on Rawls is not simply an influence on his original position argument for justice as fairness, but also an influence on his stability argument, which constitutes the second part of his argument for the two principles of justice. The importance of the stability in Rawls's overall project has been emphasized by many commentators, nevertheless, only a few recognize the importance of Rousseau's concept of amour-propre in Rawls's stability argument.[64] Rawls acknowledges the influence of Rousseau on his moral psychology, as he says, his aim in "The Sense of Justice (1963)" is to reconstruct an account of moral development in order to illustrate how "Rousseau's thesis might be true" (CP 96). What Rawls refers to by "Rousseau's thesis" is the idea that "the sense of justice is no mere moral conception formed by the understanding alone, but a true sentiment of the heart enlightened by reason, the natural outcome of our primitive affections" (CP 96).[65] We have seen that amour-propre has a reciprocal feature.

---

[64] For the importance of Rawls's stability for his overall project, see Freeman (2003; 2007), Pogge (2007), Hill (2013), Weithman (2010), Garthoff (2016).

[65] This thesis is explicit in *Emile* where Rousseau says "Our first duties are to ourselves: our primary sentiments are centered on ourselves; all our natural movements relate in the first instance to our preservation and our well-being. Thus, the first sentiment of justice does not come to us from the justice we owe but from that which is owed us; and it is again one of the mistakes of ordinary education that, speaking at first to children of their duties, never of their

According to the reciprocity thesis, those who have self-respect have the tendency to respect others. More generally, the reciprocity thesis says that those who receive benefits from others have the tendency to benefit others. This reciprocity thesis originated in Rousseau, but is developed by Rawls. In "The Sense of Justice" (1963) and Chapter VIII of *A Theory of Justice*, Rawls explains how the reciprocity of disposition is exhibited in each of the three stages of moral development. In *Restatement*, he re-emphasizes the importance of the acquisition of the sense of justice through the three stages of moral development, and argues that the reciprocity of disposition is an essential element in his reasonable moral psychology. Thus it should come as no surprise that Rawls's moral psychology has a Rousseauvian interpretation. And it is worth exploring whether other parts of the stability argument have a Rousseauvian interpretation.

In this chapter, I will first briefly introduce Rawls's moral psychology, and then I will argue that Rawls's account of moral development parallels, and is grounded in, Rousseau's thoughts on amour-propre. In particular, I argue that (1) Rawls's thesis that the sense of justice is derived from love and friendship is an illustration of Rousseau's thesis that moral sentiments are derived from natural sentiments, that (2) Rawls's explanation for how we acquire the reciprocity of disposition is grounded in Rousseau's explanation of how amour-propre can be extended into the sense of justice, and that (3) Rawls's thoughts on the principle of sympathy parallel Rousseau's thoughts on compassion. In arguing for the Rousseauvian interpretation of Rawls's moral psychology, I will also answer some of the objections against his moral psychology in a Rousseauvian way.

---

rights, one begins by telling them the opposite of what is necessary, what they cannot understand, and what cannot interest them" (E 97).

Next, I will argue that Rawls's congruence argument in *Theory* and his overlapping consensus argument in *Political Liberalism* parallel, and are grounded in, Rousseau's conception of persons. I argue that his congruence argument in *Theory* is grounded in the Kantian Rousseauvian conception of moral freedom. And I argue that his overlapping consensus argument in *Political Liberalism* is grounded in Rousseau's conception of political freedom.

Finally, I will show that Rawls's argument from the absence of special psychologies is grounded in Rousseau's concept of amour-propre. I argue that some of the objections against Rawls's treatment of economic inequality can be addressed in a Rousseauvian way.

To conclude, I argue that the three parts of the stability argument are all grounded in Rousseau's thoughts on amour-propre. Therefore, there is a Rousseauvian interpretation of Rawls's stability argument. In response to some of the objections against his stability argument, I argue that Rawls's stability argument is Rousseauvian in a deeper and better way.

1. Amour-propre and Rawls's Moral Psychology

Rawls says that his moral psychology is a psychological reconstruction which aims to illustrate how Rousseau's thesis can be true. Rousseau's thesis includes three aspects. First, he argues that moral sentiments are derived from natural sentiments. Second, he argues that amour-propre has a reciprocal feature. And third, he argues that self-love or self-esteem is the primary motivating force for human actions, not compassion. I will call the first aspect "Rousseau's thesis" narrowly construed (as Rawls calls it), the second aspect "the reciprocity thesis," and the third aspect "a realistic moral psychology." I argue that these three aspects parallel and provide the basis for Rawls's moral psychology. In this section, I will first briefly explain Rawls's moral

psychology and then I will explain the parallel relations between Rawls's moral psychology and that of Rousseau.

1.1 A Brief Introduction to Rawls's Moral Psychology

Rawls's moral psychology did not fundamentally change between his early and later works.[66] In "The Sense of Justice (1963)" he outlines the three stages of moral development through which a person acquires a sense of justice which motivates him to stabilize the basic structure of the well-ordered society. In *Theory*, this same account of moral development is restated. The role of moral psychology, however, is not limited to explaining the acquisition of a sense of justice, but also explains the congruence between the right and the thin conceptions of the good, and explains the absence of special psychology such as envy. In *Political Liberalism*, Rawls argues that there are two questions of stability, the first one is "whether people who grow up under just institutions (as the political conception defines them) acquire a normally sufficient sense of justice so that they generally comply with those institutions" (PL 141). The second question is whether the political conception of justice can be the focus of an overlapping consensus. He argues that the first question is answered by his account of moral development in *Theory*. The second question is answered by the fact that the political conception of justice is acceptable for all reasonable comprehensive views. In *Restatement*, in response to the objection that the consensus is utopian, he argues that his is a reasonable moral psychology, which includes the reciprocity of disposition that is exhibited in each stage of moral development (JF

[66] This point is made by Wenar (2005) and Estlund (1996).

196). Thus, even though the role of moral psychology shifts slightly in his later writings, the content of moral psychology remains unchanged.[67]

The content of Rawls's moral psychology includes an account of moral persons and an account of moral development to specify the account of moral persons. Rawls's moral persons are free and equal, rational and reasonable. They are free in the sense that they have the two moral powers, the capacity for a sense of justice and the capacity for a conception of the good (TJ 11; JF 21; PL 19). They are equal in the sense that they equally have these moral powers to a minimal degree to be fully cooperating members of the society (TJ 17; JF 20; PL 19). They are rational because they can choose and alter their own ends, and they seek efficient and probable means to achieve their own ends (TJ 10; JF 6-7; PL 50). They are reasonable in the sense that they are willing to honor the fair terms of cooperation, provided that others are willing to do the same. The idea of reasonableness thus is a part of the idea of reciprocity (JF 6-7; PL 40-50). Thus, to honor the two principles of justice is to honor the principle of reciprocity. They are reasonable also in the sense that they are willing to recognize the burdens of judgment and to use public reason to resolve fundamental political problems (PL 54).

The principle of reciprocity is specified in Rawls's account of moral development. His three-stage development of the morality of principles is initially stated in "The Sense of Justice" (1963) and Chapter VIII of *Theory*. The origin of moral sentiments is natural love in the infant stage. Family is the first institution of moral education. Thus Rawls says that the basic structure of the well-ordered society includes family (TJ 405). In the first stage, parents love their child, and they express their love for their child evidently through their care for him, their intention to

[67] For the roles of moral psychology in Rawls, see Krause (2008, p.35), Baldwin (2008, p.251).

do what the child's rational self-love would incline and the fulfillment of this intention. When parents' love of their child is recognized by the child, the child loves them back. This is the first psychological law. The child's returning sentiment is not immediate after he sees "the evident intention" of his parents' love. It needs to go through several steps: first, the child is assured of his self-worth when he sees "the evident intention" of his parents' love; second, the child trusts his parents and has confidence in his surrounding environment, as he experiences his parents' unconditional love; and third, in developing various skills, the child gradually acquires a sense of self-confidence, which affirms his self-esteem; and finally, the child has tightened bonds with and deep affection for his parents as he connects them with his self-development and self-esteem. The child's returning sentiment is the first appearance of the reciprocity of disposition. This stage is called "the morality of authority," for the child sees his parents as the sources of authority. And he would have a sense of guilt, "authority guilt," which is a proto-moral sentiment, if he violates their commands, provided that he has the natural sentiments of love and trust. Since the natural sentiments of love and trust are necessary for the moral sentiment of guilt, the absence of guilt indicates the absence of love and trust (TJ 425). But this stage is not a moral stage, but a proto-moral stage of moral development, for the child does not have moral sentiments in reference to moral concepts.[68]

In the proto-moral stage, the child evaluates the goodness or badness of an action according to a person's external behavior regardless of his motives. But in the second stage of moral development, as he assumes more demanding roles in more complex schemes of rights

[68] Rawls (TJ 421) argues that "it is a necessary feature of moral feelings, and part of what distinguishes them from the natural attitudes, that the person's explanation of his experience invokes a moral concept and its associated principles. His account of his feelings makes reference to an acknowledged right or wrong." Since children in the first stage of moral development do not have a clear idea of moral right or wrong, they do not have the strict sense of moral feelings.

and duties in various associations, the examples of which are family, school, games and play, and national association, he will have a more robust moral point of view. He gradually knows the standard of being a good son or daughter, a good student, a good sport and companion, and a good citizen. He learns to identify himself with others and sees things from their perspective. The development of his intellectual abilities is necessary for his acquisition of moral sensibility. But this intellectual development is not sufficient for the acquisition of moral sensibility, for a person who possesses these intellectual skills may be manipulative and exploitative. Thus, a necessary condition for moral sensibility is the recognition of others as equals.[69] By the recognition that other people are equally doing their part in the association that they belong to, he acquires the feelings of trust and friendship toward them, provided that he loves his parents back in Stage 1. If he fails to do his part, he will have a sense of guilt, "association guilt," the absence of which indicates the absence of the natural feelings of trust and friendship. In other words, the natural feelings of trust and friendship are necessary for the moral sentiment of association guilt. And so the presence of the moral sentiment of guilt indicates the presence of the natural feelings of trust and friendship. The returning sentiments of trust and friendship also need to go through several steps: the first step is to recognize that others' doing their part is beneficial to each and all in the association. Secondly, it is regarded as a form of good will that others are doing their part and they want to, and this arouses feelings of trust and friendship. Thirdly, everyone who participates in the same activity reciprocally does his or her part in the association.

In a certain form of the morality of association, once a person has the recognition that his fellow citizens willingly contribute to the cooperative scheme of a well-ordered society, he will

---

[69] As we have seen that in the first stage, children are affirmed of self-respect through their parents' manifest love and the development of their own intellectual skills. The sense of self-respect inclines them to respect others and recognize them as equals.

have friendly feelings and trust toward them, and he will reciprocally do his share. He will have a sense of guilt if he fails to contribute to social cooperation, provided that he has feelings of trust and friendship toward them. But since the natural feelings of trust and friendship are necessary for the moral sentiment of guilt, his moral sentiments are contingent upon his natural feelings of trust and friendship, without which, he would not have "association guilt." In the final stage of moral development, the morality of principles, a person acquires a sense of justice if he recognizes that the just institutions benefit him and those whom he cares for, provided he has acquired feelings of love and friendship in the first two stages. Unlike the second stage, now his moral sentiments are not contingent upon his natural feelings toward his fellow citizens, but upon the acceptance of the principles of justice. For while a citizen is attached to several other citizens, nobody is attached to all. And so citizens cannot be bound with each through the ties of friendship; rather, they are bound by the allegiance to the principle of justice. Note that the acquisition of love and friendship in the first two stages is still necessary for the sense of justice. But once a person has the sense of justice, he will act upon the principles of justice no matter if he has feelings of love or friendship toward others or not. If he violates the principles of justice, he will have a sense of guilt, "principle guilt," regardless of the contingencies of the world. But if he at the same time has fellow feelings toward his countrymen, his moral sentiments will be intensified. Rawls thus says that the sense of justice and the love of humankind are two forms of the morality of principles. But he argues that the later is supererogatory while the former is a natural duty. He concludes that justice as fairness has a more realistic moral psychology than other conceptions of justice. It is more realistic than intuitionism because intuitionism fails to explain the sense of justice in a psychologically understandable way (TJ 418). And even though utilitarianism and perfectionism explain the sense of justice in a natural way, he suggests that

they ground justice in the love of humankind, a supererogatory duty.[70] What makes Rawls's moral psychology realistic is that it recognizes the primary motivating force of self-respect, which inclines a person to respect others. The principle of reciprocity is neither egoistic nor altruistic.

To sum up, the content of Rawls's moral psychology has three themes. First, the sense of justice is derived from love and friendship. In other words, to acquire the sense of justice in the final stage, a person must love his parents back in the first stage and have friendly feelings toward his associates in the second stage. Second, the reciprocity of disposition is exhibited in each stage of moral development. In the first stage, once the child recognizes his parents' love with evident intention, he will love his parents back. In the second stage, given that the child loves his parents back in the first stage, once he sees that his associates do their parts in the associations they belong to, he will have feelings of trust and friendship toward them. In the third stage, given that he has returning sentiments of love and friendship in the first two stages, once he sees that the just institutions are beneficial to him and those around him, he will acquire a sense of justice. Third, Rawls's moral psychology is a realistic moral psychology. It is more realistic than utilitarianism and perfectionism because for him the love of humankind which is their psychological basis is a supererogatory duty.

---

[70] Rawls connects utilitarianism with the principle of sympathy or the love of humankind in several places, for example, see section 5, section 30 in *Theory*. For Rawls, both perfectionism and utilitarianism may sacrifice freedom for the sake of the greater good (perfection or utility), and so their moral psychologies are similarly unrealistic.

1.2 The Parallel Relations Between Rousseau's Concept of Amour-propre and Rawls's Moral Psychology

I argue that Rawls's moral psychology parallels, and is grounded in, Rousseau's moral education in *Emile*.[71] First, Rawls's thesis that the sense of justice is derived from love and friendship is an illustration of Rousseau's thesis that moral sentiments are derived from natural feelings. Second, Rawls's explanation of how we acquire the reciprocity of disposition is grounded in Rousseau's explanation of how amour-propre can be extended to the sense of justice. And third, Rawls's thoughts on the principle of sympathy or the love of humankind parallel Rousseau's realistic moral psychology.

1.2.1 Rousseau's Thesis

Rawls acknowledges that his moral psychology is an illustration of how Rousseau's thesis might be true (CP 96). He says that Rousseau's thesis is "the sense of justice is no mere moral conception formed by the understanding alone, but a true sentiment of the heart enlightened by reason, the natural outcome of our primitive affections" (CP 96). His account of moral development tries to show that a person's sense of justice is not simply formed by the understanding or recognition that the just institutions are beneficial to him and those around him, but also the outcome of the natural feelings of love and friendship. It is true that in the final stage of moral development, the sense of justice is not contingent on the natural feelings. But without the acquisition of the natural feelings of love and friendship in the first two stages, it is

[71] Samuel Freeman (2007, p.19) points out that Rawls largely agrees with Rousseau's description of human nature and that Rawls's moral psychology in largely influenced by Rousseau's moral education in *Emile*.

impossible to acquire the sense of justice. Or we may say, the acquisition of the natural feelings of love and friendship is a necessary condition for the acquisition of the sense of justice.

John Deigh (1996) argues that Rawls fails to explain how the sense of justice is derived from love and friendship. This failure, he argues, comes from Rawls's confusion between guilt and remorse. He argues that the moral feeling, which is necessarily connected with the natural feelings of love and friendship, is remorse, and that Rawls mistakenly confuses this feeling with guilt. He argues that this confusion leads Rawls to claim that a sense of guilt connects the sense of justice with the natural feelings of love and friendship. But since guilt is not necessarily connected with the natural feelings of love and friendship, he argues, the sense of justice is not derived from love and friendship. He says:

> To explain how the child comes to love its parents is not yet to explain how it comes to respect them. The former establishes that the child identifies with its parents but not that it accepts their authority, and acceptance of authority is the facet of moral development Rawls means to explain in giving his account of the first stage. (Deigh, 1996, p.57)

Deigh correctly points out that love is not sufficient for "authority guilt." It is not sufficient to explain how children come to respect parents' authority. It is also not sufficient to explain how citizens in the well-ordered society of justice as fairness acquire the sense of justice. Rawls acknowledges that love is merely a necessary but not sufficient condition for "authority guilt." He writes:

> In the light of this sketch of the development of the morality of authority, it seems that the conditions favoring its being learned by the child are these. First, the parents must love the child and be worthy objects of his admiration. In this

> way they arouse in him a sense of his own value and the desire to become the sort of person that they are. Secondly, they must enunciate clear and intelligible (and of course justifiable) rules adapted to the child's level of comprehension. In addition they should set out the reasons for these injunctions so far as these can be understood, and they must also follow these precepts insofar as they apply to them as well. (TJ 408)

Thus, for Rawls, love is merely a necessary but not sufficient condition for moral sentiments. For a child to acquire "authority guilt," he not only needs to have parental love but also understand the rules that are set up and followed by his parents. But Deigh not only denies that love is sufficient for moral sentiments, he also denies that love is necessary for moral sentiments. He endorses Freud's pessimistic view of moral education and explains guilt in terms of fear. He writes:

> At this stage, fear is the emotion that they in their role as authorities evoke in the child. Then, as their authority becomes internal, as the child comes to see itself as part of a moral order over which its parents preside, its fear of their power is transformed into respect for their authority. Correspondingly, its fear of punishment — or of whatever evil the child senses will be the consequence of disobedience — becomes a liability to guilt. (Deigh, 1996, p.62)

Here Deigh argues that a person has feeling of guilt only if he or she is afraid of punishment. But Deigh confuses the acceptance of authority with the acquisition of moral sentiment. To explain why a person accepts parental (or anyone's) authority does not yet explain why this person has the moral sentiment of guilt when he violates his parents' (or anyone's) demands. In fact, a person may accept his parents' (or anyone's) authority, if they have a gun to

his head, but does not have any feelings of guilt when he violates their demands. In any case, Deigh does not explain why and how a person internalizes an external authority. Fear of punishment may explain why a person accepts an external authority but it does not explain why and how this person internalizes such authority.[72] Rawls's and Rousseau's moral psychologies fill in the blank. Rawls agrees with Rousseau that love is the origin of moral sentiments. Without the natural sentiment of love, the moral sentiment of guilt is inexplicable. Fear cannot explain the moral sentiment of guilt. A person may do as someone commands out of fear but he will not have the moral sentiments of guilt if he violates this person's commands. Rousseau warns us about the consequences of imposing punishment on children:

> Firstly, by imposing on them a duty they do not feel, you set them against your tyranny and turn them away from loving you. Secondly, you teach them to become dissemblers, fakers, and liars in order to extort rewards or escape punishments. Finally, by accustoming them always to cover a secret motive with an apparent motive, you yourselves give them the means of deceiving you ceaselessly, of depriving you of the knowledge of their true character, and of fobbing you and others off with vain words when the occasion serves. (E 91)

Rousseau writes:

> It is quite strange that since people first became involved with raising children, no instruments for guiding them has been imagined other than emulation, jealousy, envy, vanity, avidity, and *vile fear* — all the most dangerous passions, the quickest to ferment and the most appropriate to corrupt the soul, even before the

[72] In certain cases, fear of punishment or power cannot explain why a person accepts an external authority. Even children have self-esteem, and once their self-esteem is hurt by the punishments imposed on them, they will not obey the external authority.

> body has been formed. With each lesson that one wants to put into their heads before its proper time, a vice is planted in the depth of their heart. (E 92)[73]

Thus, for Rousseau, fear is not the appropriate means of making a child a good person. On the contrary, fear is "most appropriate to corrupt the soul." A child who is educated in this way would be either servile or arrogant. This child, in Rousseau's terms, would have an inflamed amour-propre. We have seen in the first chapter that the inflamed amour-propre is the root of all evils. Thus, if in the early stage of his moral development, a person's amour-propre is inflamed, it will be impossible for him to acquire the sense of justice in the later stages. Rousseau argues that the sense of justice is derived from primitive sentiments. He writes:

> Our first duties are to ourselves; our primary sentiments are centered on ourselves; all our natural movements relate in the first instance to our preservation and our well-being. Thus, the first sentiment of justice does not come to us from the justice we owe but from that which is owed us. (E 97)

For Rousseau, since self-love is reciprocal (E 214), a person who is affirmed of his self-love by his parents' manifest love tends to love his parents back. If this person "sees that everything approaching him is inclined to assist him" (E 213), he will have friendly feelings toward those around him. Eventually, he will acquire the sense of justice.

To sum up, Rawls's moral psychology is an illustration of Rousseau's thesis that moral sentiments are derived from natural sentiments, or, in particular, the sense of justice is derived from love and friendship. Deigh argues that Rawls's moral psychology mistakenly grounds justice in love. He argues that the sense of justice is derived from fear. But as we have seen, in

---

[73] Italics added.

Rousseau, fear is not the appropriate way to nurture the sense of justice. On the contrary, it gives rise to an inflamed amour-propre, which, as we have seen, is the root of all evils.

1.2.2 The Reciprocity Thesis

In the previous chapter, we saw that Rawls's concept of reciprocity is derived from Rousseau's concept of amour-propre, more specifically, the reciprocal feature of amour-propre. Rousseau suggests that our natural and healthy form of amour-propre makes a rightful claim on our conduct, a claim to grant others equal standing with us provided they do the same in return. And so those who have self-respect are more likely to respect others. This explains why the natural and healthy form of amour-propre is the psychological foundation of a society of mutual respect and self-esteem for both Rousseau and Rawls. But it remains to be seen how citizens in the well-ordered society of justice as fairness acquire the reciprocity of disposition which inclines them to accept the principle of reciprocity or the principles of justice.[74] Rawls's moral psychology provides an answer to this question. Rawls says that "In the account in *Theory* of the three-stage development of the morality of principles (as it is called there), the psychological laws for each stage exhibit this reciprocity of disposition" (JF 196). The reciprocity of disposition is a "tendency to answer in kind, to answer others' being fair to us with our being fair to them, and the like" (JF 196). In the first stage of moral development, children acquire the reciprocity of disposition for the first time. They are disposed to love their parents back once they recognize that their parents love and want to love him. In the second stage, as they play

[74] In "Justice as Reciprocity" (1971), Rawls says that "The principle of reciprocity requires of a practice that it satisfy those principles which the persons who participate in it could reasonably propose for mutual acceptance under the circumstances and conditions of the hypothetical account" (CP 208). And he argues that the principles of justice satisfy the principle of reciprocity while utilitarianism does not.

more complex roles in the society, they reciprocally contribute to the cooperative arrangement of the association once they recognize that others willingly contribute to it. And in Stage 3, they willingly comply with the two principles of justice once they recognize that the just institutions are beneficial to them and those around them. It is important to note that the reciprocity of disposition is acquired for the first time when the children are disposed to love their parents back on the condition that their parents love them "with evident intention." As Rawls says:

> The parents, we may suppose, love the child and in time the child comes to love and to trust his parents. How does this change in the child come about? To answer this question I assume the following psychological principle: the child comes to love the parents only if they manifestly first love him. (TJ 406)

Rawls notes that this psychological principle is drawn from Rousseau's *Emile* (TJ 406, fn.9). In explaining the second psychological law that people acquire feelings of trust and friendship toward their associates on the condition that their associates do their part "with evident intention," He notes that "the idea of evident intention as used here is from Rousseau's *Emile*" (JF 196, fn.16). In the *Emile*, Rousseau says:

> A child's first sentiment is to love himself; and the second, which derives from the first, is to love those who come near him, for in the state of weakness that he is in, he does not recognize anyone except by the assistance and cares he receives. At first the attachment he has for his nurse and his governess is only habit. He seeks them because he needs them and is well-off in having them; it is recognition rather than benevolence. He needs much time to understand that not only are they useful to him but they want to be; and it is then he begins to love them. (E 213)

Rousseau points out that children are disposed to return love for love if the caregivers manifestly love them. This is the first step of the extension of self-love. Human beings naturally have self-love and they naturally love those who manifestly love them. Love is extended from themselves to those around them (provided that they love them with evident intention). Parental love planted the seed of humanity in Emile so that he could become a good person in his later life. Once Emile acquired the returning sentiment of love, in reaching the adulthood, after he became interested in his surroundings (or after he began to compare himself with others around him), he acquired the first sentiment of friendship toward others. In this step, self-love becomes relative; amour de soi turns into amour-propre. But Emile is brought up "(1) To have a sense of the true relations of man, with respect to the species as well as the individual. (2) To order all the affections of the soul according to these relations" (E 219). Thus Emile's amour-propre is not inflamed. He recognizes that human beings have equal status and that they are not in the relations of subjugation and domination, alienation and exploitation. By recognizing others as equals, Emile's amour-propre extends to the love of human species, or the love of justice. Now, for Rousseau there are two sources of the sense of justice: the recognition of human beings as equals and the natural feelings of love and friendship. And the natural feelings of love and friendship are aroused in his heart because of his recognition of the manifest love of his parents and those around him. As Rousseau says, "A child is therefore naturally inclined to benevolence, because he sees that everything approaching him is inclined to assist him; and from this observation he gets the habit of a sentiment favorable to his species" (E 213).

Thus, not only is Rawls's reciprocity thesis grounded in Rousseau's reciprocal feature of amour-propre, but Rawls's explanation of how citizens in the well-ordered society of justice as fairness acquire the reciprocity of disposition which inclines them to accept the principle of

reciprocity or the principles of justice is grounded in Rousseau's explanation of how amour-propre can be extended to the sense of justice.

1.2.3 A Realistic Moral Psychology

Rawls suggests that justice as fairness has a more realistic moral psychology than other conceptions of justice, such as utilitarianism, perfectionism and intuitionism. The flaw of intuitionism is that it fails to describe the sense of justice as "psychologically understandable." Rawls says that for intuitionism, the sense of justice is simply a desire for what is right but the right is unanalyzable. And thus intuitionism fails to explain the sense of justice in an understandable way. Rawls argues that utilitarianism and perfectionism can interpret the sense of justice in a natural way, and, thus, they can describe the sense of justice as "psychologically understandable" (TJ 417). But he suggests that the love of humankind, which is the psychological foundation for utilitarianism and perfectionism, is supererogatory for human beings. And thus utilitarianism's (and perfectionism's) moral psychology is the psychology of the saints (and the psychology of heroes), not a realistic moral psychology for ordinary human beings. Thus, for Rawls, justice as fairness provides a more realistic moral psychology than other conceptions of justice.

Rawls's realistic moral psychology parallels Rousseau's thoughts on amour-propre. In particular, Rousseau argues that amour-propre, rather than compassion, is the primary motivating force for human actions. As we have seen, in the second *Discourse*, Rousseau suggests that even though amour de soi and compassion are the two principles of action in the state of nature, amour de soi, rather than compassion, is the primary motivating force. He also says that in the civil state, compassion is too weak to motivate a person to act compared with amour-propre. We have

also seen that in *Emile*, Rousseau suggests that amour-propre has a natural and healthy form if guided by reason, and that it is the starting point for justice. He argues that compassion plays a secondary, supporting role in the transformation of amour-propre, but it is not the starting point for justice. We have also seen that in the *Social Contract* and other of Rousseau's political works, he suggests that once amour-propre is guided by the general will (public reason), it will be the primary motivating force to form and retain the just society. Compassion, thus, is not the primary motivating force to stabilize the just society.

Michael Pritchard (1977) argues that Rawls's moral psychology faces a dilemma: either his principle of equal respect is simply based on the principle of reciprocity, as he claims it to be, in which case there is a gap between the principle of equal respect and the principle of reciprocity; or, his principle of equal respect also relies on the principle of sympathy, as utilitarianism does, in which case his objection against utilitarianism fails. To explain the first horn of the dilemma, Pritchard argues that the principle of reciprocity cannot generate the sense of justice. In other words, he argues that moved by the principle of reciprocity, people will not have the sense of justice toward those who cannot contribute to the social cooperation of the well-ordered society.[75] For example, he says that the disabled are incapable of contributing to the cooperative scheme of the well-ordered society, and, thus, according to Rawls they do not deserve equal respect.[76] To explain the second horn of the dilemma, he argues that the principle of reciprocity

---

[75] For similar objections, see Brian Barry (1995, p.31), Allan Buchanan (1990, pp.228-30), G.A Cohen (1995, p.224), Martha Nussbaum (2006, p.123).

[76] Prichard (1977) says that "suppose that reliable methods were developed for determining which among newborns would create the greatest problems for the society-e.g., those with serious intellectual or physical disabilities, those who will contract crippling or disabling diseases relatively early in life, or those who will end up being the dissenters in the society and try to bring down its institutions. If the society adopted the practice of removing such newborns from the society (either by infanticide or by sending them to less well-ordered societies), this would seem to be consistent with principles of reciprocity, assuming the newborns had made no significant contribution to the society" (p.65). But on the one hand, Rawls might say that those with serious intellectual or physical disabilities can

cannot generate the principle of sympathy. He argues that the principle of reciprocity cannot generate the love of humankind, which, he maintains, is the rightful basis of the just society because it explains the equal respect for each and all.

But Pritchard fails to see the distinction between love and respect. Rawls points out that the love of humankind is supererogatory for ordinary human beings. As Rawls says, "This morality [the love of humankind] is not one for ordinary persons, and its peculiar virtues are those of benevolence, a heightened sensitivity to the feelings and wants of others, and a proper humility and unconcern with self" (TJ 419). Thus it is unrealistic to think that all human beings have "a heightened sensitivity." Furthermore, the love of humankind is not necessary for a society of mutual equal respect. What is needed is the reciprocity of disposition.

Pritchard also fails to fully understand Rawls's principle of reciprocity. He fails to understand why the reciprocity of disposition will generate mutual equal respect. He thinks that the principle of reciprocity is the morality of exchange. On this reading, since those who are disabled do not contribute to the cooperative scheme of the well-ordered society, they are not the objects of equal respect. But Rawls never merely understands the reciprocity of disposition in this way. Indeed, in the first stage of moral development, the morality of authority may be the morality of exchange because the returning sentiments of love are contingent on the love they receive from their parents. But in the second stage, the morality of association may not be the morality of exchange; for we may still do our parts in the associations we belong to, even though we know that there are free-riders in the associations. And in the final stage of moral development, the morality of principles, the sense of justice is not contingent on the benefits they

---

make significant contribution to the society; on the other hand, he might say that even though they cannot make contribution to the society due to the misfortune in the natural lottery, they still deserve equal respect.

receive from others. Rawls says: "Once a morality of principles is accepted, however, moral attitudes are no longer connected solely with the well-being and approval of particular individuals and groups, but are shaped by a conception of right chosen irrespective of these contingencies" (TJ 416). In the final stage of moral development, recognizing that the just institutions are beneficial to them and those around them, citizens will have the sense of justice, provided that they have feelings of love and friendship in the first two stages. Their sense of justice is not toward those they love and befriend, for, if so, they will still be in the first two stages of moral development. The mark of their being in the final stage is that their sense of justice is not dependent on the benefits they receive from particular persons. Pritchard is wrong, therefore, to say that for Rawls those who have a sense of justice will not have equal respect to those who cannot contribute to the cooperative scheme of the well-ordered society.

But perhaps Pritchard's question is how it is possible that those who have a sense of justice will have equal respect with those who cannot contribute to the cooperative scheme of the well-ordered society. In other words, his question might be about how it is possible that human beings can proceed into the final stage of moral development. The empirical psychologies of Piaget and Kohlberg are also subject to the same question. But, unlike their empirical psychology, Rawls's moral psychology is not a science of human nature (PL 86). His reasonable moral psychology is compatible with human nature, and even more so than the moral psychologies of utilitarianism and perfectionism. Rawls does not need to show that his moral psychology is the most realistic one; he only needs to show that it is more realistic than the moral psychologies of other conceptions of justice. Furthermore, if Rawls's moral psychology is plausible, it will provide support for Piaget's and Kohlberg's empirical psychologies and refute the objections against their moral psychologies. It is simply too early to say that it is impossible that human beings can

develop into the final stage of moral development. After all, we do not live in the just society that Rawls depicts as a realistic utopia. And that most of the people in an unjust society cannot develop into the final stage of moral development says nothing about the moral psychology of those who live in the just society. For Rawls and Rousseau, institutional factors have significant impact on what human beings are like, and so we can reasonably expect that in the just society people will have the sense of justice in the final stage of moral development.

To conclude, I have argued that there is a Rousseauvian interpretation of Rawls's moral psychology which constitutes the first part of his stability argument, and that on this interpretation, some of the objections against his moral psychology can be answered. In response to these objections, I hope that I have shown that Rawls's moral psychology is Rousseauvian in a deeper and better way.

2. Amour-propre, the Congruence Argument and the Overlapping Consensus Argument

Rawls's moral psychology has shown us how the sense of justice can be developed through the three stages of moral development. We have seen that his moral psychology parallels, and is grounded in, Rousseau's thoughts on amour-propre and that the objections against his moral psychology can be answered on my Rousseauvian interpretation. But it remains to be seen whether the congruence argument in *Theory* and the overlapping consensus argument in *Political Liberalism* have a Rousseauvian interpretation. I argue that these arguments are grounded in Rousseau's conception of persons. The congruence argument in *Theory* is grounded in Rousseau's conception of moral freedom. This conception of moral freedom is developed by Kant and endorsed by Rawls in *Theory* as the moral basis of the justification for the two

principles of justice. The overlapping consensus argument in *Political Liberalism* is grounded in Rousseau's conception of political freedom. The citizens who guide amour-propre with public reason are politically autonomous. This ideal of political autonomy is endorsed by the later Rawls to explain why the political conception of justice is acceptable for all reasonable comprehensive views.

2.1 Amour-propre and the Congruence Argument in *Theory*

The congruence argument in *Theory* argues that acting from the sense of justice is congruent with the thin conceptions of the good. This argument is called the Kantian Congruence argument because it starts with a Kantian premise that persons are free and equal rational beings (cf. Freeman 2007, p.274; Weithman 2010, p.206). The fundamental preference of the moral persons is the desire to express themselves as free and equal rational beings (TJ 491). "To express one's nature as a being of a particular kind is to act on the principles that would be chosen if this nature were the decisive determining element" (TJ 222). Since the Original Position is a situation in which our nature is a decisive determining element (TJ 221), to express one's nature as a free and equal rational being is to act on the principles that would be chosen in the Original Position. Therefore, to express oneself as a free and equal rational being is to act on the principles that would be chosen in the Original Position. Therefore, the desire to express oneself as a free and equal rational being is the same desire as acting from the sense of justice (TJ 501). According to the finality criterion, to satisfy the desire to act on the principles that would be chosen in the Original Position (or the sense of justice), one must take that desire as supremely regulative (TJ 503). Therefore, the desire to express oneself as a free and equal rational being is the same as the desire to treat the sense of justice as supremely regulative of

other desires. The congruence conclusion is established: members of a well-ordered society judge, from within the thin theory of the good, that the sense of justice is supremely regulative of their other desires.

The Kantian congruence argument is partly Rousseauvian. The parallel relations between Rousseau's conception of moral freedom and Kant's can be seen from the following passages:

> To the preceding one might add to the credit of the civil state moral freedom, which alone makes man truly the master of himself; for the impulsion of mere appetite is slavery, and obedience to the law one has prescribed to oneself is freedom. (SC 54).
>
> *Will* is a kind of causality of living beings insofar as they are rational, and *freedom* would be that property of such causality that it can be efficient independently of alien causes *determining* it… The preceding definition of freedom is *negative*, and therefore unfruitful for insight into its essence; but there flows from it a *positive* conception of freedom, which is so much the richer and more fruitful… The will is in all its action a law to itself, indicates only the principle, to act on no other maxim than that which can also have as object itself as a universal law. (Kant, 1996, p.94)[77]

Rousseau's conception of moral freedom has both a negative sense and a positive sense. The negative sense of moral freedom is the absence of animal appetites. The positive sense of moral freedom is the obedience to the law one gives oneself. This conception of moral freedom is endorsed and developed by Kant to explain the autonomy of the will. For Kant, moral freedom

[77] Italics original.

also has both a negative sense and a positive sense. The negative sense of moral freedom is the independence from alien causes. The positive sense of moral freedom is to act on the universal law. The parallel relations between Rousseau's conception of moral freedom and that of Kant thus show that the latter's conception of moral freedom is grounded in the former's conception of moral freedom. As is pointed out by Rawls, "Kant's main aim is to deepen and to justify Rousseau's idea that liberty is acting in accordance with a law that we give to ourselves. And this leads not to a morality of austere command but to an ethic of mutual respect and self-esteem" (TJ 225). The Kantian and Rousseauvian moral freedom provides the moral basis for a society of mutual respect and self-esteem. For the Kantian and Rousseauvian persons see themselves as free and equal moral persons by nature. For them, acting justly is an expression of their moral nature, and their fundamental preference. And so, acting justly is congruent with their thin conception of the good. What underlies the Kantian Rousseauvian congruence argument is the Kantian Rousseauvian conception of moral persons. The reason why the Kantian Rousseauvian persons act on and from the two principles of justice is that they see acting from the principles of justice as an expression of their personal integrity, a requirement of moral autonomy.

But it is unreasonable to expect that all reasonable persons in the well-ordered society of justice as fairness are Kantian Rousseauvian persons. Many of them have very different comprehensive views. Christians may take honoring God as the expression of their nature, and utilitarians may take acting from the principle of utility as the expression of their nature. For those who only have a partial comprehensive view, they may have plural ends, and they may not order these ends by a simple principle. It is, thus, also unreasonable to expect them to take acting justly as the expression of their nature. In a society regulated by a comprehensive doctrine, people who have a comprehensive view may not respect other comprehensive views. Thus, as

Rawls later suggests, the Kantian and Rousseauvian moral freedom is not the correct path toward the society of mutual respect and self-esteem.

Rousseau never takes his conception of moral freedom to be essential for his social contract theory, as he says: "But I have already said too much on this topic, and the philosophical meaning of the word *freedom* is not my subject here" (SC 54). What is essential for Rousseau's social contract theory is his conception of political freedom, the conformation to the general will. As we shall see later, this conception of political freedom is what grounds Rawls's congruence argument in *Political Liberalism*.

2.2 Amour-propre and the Overlapping Consensus Argument in *Political Liberalism*

In *Political Liberalism*, Rawls says that the stability argument in *Theory* is inconsistent with his theory of justice as a whole (PL xvi). For the citizens endorse the two principles of justice on the basis of a comprehensive philosophical view, that is, comprehensive liberalism. He realizes that, based on the fact of reasonable pluralism, any conception of justice, which is grounded in a comprehensive doctrine, cannot be stable for the right reason. For a comprehensive doctrine cannot but be in conflict with other comprehensive doctrines. He argues that his political conception of justice is the focus of overlapping consensus and that it can be stable for the right reason, for this political conception of justice is acceptable for all reasonable comprehensive views. Thus, for example, Kantians, Christians, utilitarians and those who have partial comprehensive views all have different views as to what expresses their nature, but as long as they appeal to public reason, namely the reason that cannot be reasonably rejected by others, to solve fundamental public issues, there will be an overlapping consensus between their

comprehensive views and the political conception of justice. Rawls argues that if the overlapping consensus argument sustains, his stability argument is defensible.

I argue that Rawls's overlapping consensus argument in *Political Liberalism* is grounded in Rousseau's conception of political freedom. For Rawls's conception of public reason parallels Rousseau's conception of the general will. As we have seen, the educative purpose of Emile is to make Emile an autonomous person who guides his amour-propre with reason. In the *Social Contract*, Rousseau makes it clear that one's particular will must conform to the general will and that it is through the general will that one knows whether his opinion is true or false. Thus, for Rousseau, citizens must appeal to the general will to resolve fundamental public issues. And it shall be clear that, for him, citizens willingly comply with the general will because the general will is a part of the will of their own and because it is through submitting themselves to the general will that they remain free in the civil state. In other words, citizens willingly comply with the just institutions in part because the just institutions satisfy the natural and healthy form of amour-propre, the desire for freedom and equality. Therefore, political freedom provides another path, the correct path as the later Rawls says, toward a society of mutual respect and self-esteem.

Parallel to Rousseau's conception of political freedom, Rawls argues that political freedom is one of the primary goods essential for securing the social bases of self-respect. Rawls argues that the social primary goods, rights and liberties, powers and opportunities, income and wealth, and the social bases of self-respect, are necessary for the pursuit of any rational plan of life. They are also necessary for the exercise and development of the two moral powers. Thus, for all free and equal, rational and reasonable persons in the well-ordered society of justice as fairness, irrespective of whether they are Kantians, Christians, or utilitarians, they will all value the fair equal distribution of these social primary goods. Thus, once they recognize that the well-

ordered society of justice as fairness ensures the fair equal distribution of these social primary goods, they will willingly comply with the rules and the laws in the well-ordered society of justice as fairness. In particular, self-respect is the most important primary good. Once these persons see that the public affirmation of the two principles of justice and their lexical order ensures their self-respect, they will willingly comply with the just intuitions and their rules and laws.[78] As Rawls says, "moved by amour-propre we are ready to accept and to act on a principle of reciprocity whenever our culture makes it available and intelligible to us, and society's basic arrangements establish our secure and equal standing alone with others" (LHPP 199). In general, as beings who are moved by amour de soi and amour-propre, once we recognize that the just institutions ensure the equal distribution of liberties and rights, and the fair distribution of other necessary means of our conceptions of the good, we will willingly comply with the two principles of justice.[79] This brings about the stability of Rawls's well-ordered society of justice as fairness.[80]

The stability of Rawls's well-ordered society allows a small number of unreasonable people and a certain range of indeterminate fundamental political questions in the well-ordered society

---

[78] Nussbaum (2006, p.225) points out that even though political liberals and communitarians have very different opinions on many matters, they will all think that mutual-respect and reciprocity are extremely important social goods.

[79] The psychological fact that we are moved by amour de soi and amour-propre is not a comprehensive view.

[80] Gerald Gaus (2013, p.244) maintains that Rawls's argument that citizens in the well-ordered society of justice as fairness will willingly comply with the two principles of justice because of the fair equal distribution of social primary goods appears only in *A Theory of Justice*. He argues that in *Political Liberalism* Rawls merely tries to show that the political conception of justice is not in conflict with all reasonable comprehensive views. But there is no reason to think that Rawls has abandoned his theory of primary goods in *Political Liberalism*, as Rawls says: "at the basis of the parties' reliance on primary goods is their recognition that these goods are essential all-purpose means to realize the higher-order interests connected with citizens' moral powers and their determinate conceptions of the good (so far as the restrictions on information permit the parties to know this). The parties are trying to guarantee the political and social conditions for citizens to pursue their good and to exercise the moral power that characterize them as free and equal" (PL 76).

in justice as fairness.[81] Given that we can expect that most citizens in the well-ordered society in justice as fairness have at least acquired friendly feeling toward their fellow citizens, if not the sense of justice, it is reasonable to think that there will not be a large number of unreasonable people and a variety of indeterminate fundamental political questions.[82]

3. Amour-propre and Rawls's Argument from the Absence of Special psychology

We have seen that the two questions of stability have been answered by Rawls's moral psychology and his overlapping consensus argument. We have also seen that Rousseau's concept of amour-propre plays a significant role in the two parts of the stability argument. Thus we are ready to say that there is a Rousseauvian interpretation of Rawls's stability argument. But it remains to be seen whether citizens in the well-ordered society of justice as fairness would be free from special psychologies, in particular, social envy, spite, a will to dominate or a tendency to submit. In *Restatement*, Rawls says that if we have a sufficiently strong and effective sense of justice, and if we have the recognition that the well-ordered society of justice as fairness is good for us, we are unlikely to be moved by these special psychologies (JF 181, 202). In *Theory*, on the other hand, he has a subtler and more nuanced argument to explain why the well-ordered society in justice as fairness will not generate these special psychologies. In section 29, he argues

---

[81] Michael Huemer (1996) argues that Christian fundamentalists are not reasonable and that they will not agree on the political conception of justice. Robert Talisse (2003) argues that Millian utilitarianism are not reasonable and that they will not agree on the political conception of justice. But neither Christian fundamentalists nor Millian utilitarians are necessarily unreasonable. If they appeal to public reason in the political sphere, they can be reasonable despite the fact that they give priority to the Christian doctrines or the principles of utility in ordering their goods in their private lives. For examples of two sorts of indeterminate fundamental political questions, see David Reidy (2000).

[82] For a detailed theory of moral education from a Rawlsian perspective, see Victoria Costa (2010).

that self-respect is mutually supportive and that self-contempt is mutually disadvantageous. In section 81, he argues that a robust sense of self-respect consists of a kind of psychological immunity toward special psychologies such as envy.[83] He says that envy is a collective disadvantageous feeling ready to worsen the situation of each party, the one who envies and the one who is envied.[84] He, thus, regards envy as a destabilizing force for the well-ordered society of justice as fairness. He argues that there are three conditions that encourage envy. These three conditions are an inadequate sense of self-worth, the large visible differences between oneself and others, and insufficient constructive alternatives, namely, associations or communities, to opposing the favored condition of the more advantaged in the existing social arrangement. He argues that since the well-ordered society in justice as fairness ensures one's sense of self-worth through the equal distribution of liberties and rights, reduces social inequalities that are not to the greatest benefit of the least advantaged people, and allows various associations and communities to oppose the favored condition of the more advantaged, such a society would not generate envy to undermine its stability.

Careful readers will discover that these three conditions are all closely connected with the social bases of self-respect. In section 29, Rawls argues that the two principles of justice support men's self-esteem. He argues that for the equal distribution of basic liberties and rights and the fair distribution of other means of exercising and development the two moral powers and the conceptions of the good ensure the social bases of self-respect. As he says: "For by arranging inequalities for reciprocal advantage and by abstaining form the exploitation of the contingencies

---

[83] Rawls mainly discusses the problem of envy in *Theory*, but he says that the problem of other special psychologies can be treated in the same way as envy (TJ 474).

[84] We may say that other special psychologies are also mutually disadvantageous for both parties. Rawls explicitly says that self-contempt is mutually disadvantages in section 29. This may also be true for spite, the will for domination and the tendency for submission.

of nature and social circumstance within a framework of equal liberties, persons express their respect for one another in the very constitution of their society" (TJ 155). In section 82, he argues that the equal distribution of fundamental rights and liberties is the legitimate social base of respect and that the need for equal status is met by the public affirmation of equal citizenship and "the full and diverse internal life of the many free communities of interests that the equal liberties allow" (TJ 477). Since having a secure sense of amour-propre is the main reason why citizens in the well-ordered society are free from special psychologies such as envy, Rawls's argument from the absence of special psychologies is grounded in Rousseau's concept of amour-propre.

Some commentators argue, against Rawls, that the equal distribution of basic liberties and rights is not sufficient for a secure sense of self-respect.[85] They argue that economic equality might also be important for the secure sense of self-respect. For economic inequality is also a major contributor to social envy. As we have seen, Rawls also thinks that large economic inequalities caused by the social structure of a society can be a source of envy and a destabilizing force for the society. But he does not think that the distribution of wealth and income should be the legitimate base of self-respect because grounding self-respect in social-economic status necessarily arouses the desire for domination or the tendency for submission. As Rousseau points out, the true value of human beings lies in freedom, and equality as a means of freedom. Unless people recognize that their true value lies in their equal political status, not socio-economic

[85] For example, Bercuson (2014, p.67) says that "We ought not accept (at face value) the claim that institutional equality can prevent the envy associated with distributive inequality. Although in some instances of public debate, we can imagine economic inequality receding into the background (and not affecting constitutional or policy outcomes), in other instances of political deliberation the self-confidence or influence of involved parties is directly tied to their distributive shares, regardless of norms of moral desert." Neuhouser (2014, p.223) also makes a similar point.

status, they will not have a natural and healthy form of amour-propre, a desire for freedom and equality. The public recognition of the two principles of justice and their lexical order is thus necessary for people to have a sufficient sense of self-respect. This does not mean that economic inequalities can be ignored in the well-ordered society of justice as fairness. For large economic inequalities caused by the basic structure of the society will have substantial impact on the alleged equal basic liberties and rights that citizens possess. Thus, Rawls argues that social and economic inequalities are be to arranged in such a way that they are to the greatest advantage of the least advantageous people. Thus, for Rawls, social and economic inequalities are only permissible when these inequalities are better for the less advantaged people than social and economic equality. Under this circumstance, if the less advantaged people are still envious of the more advantaged people, we may say that such envy is unreasonable and should not be considered in a theory of justice.

To conclude, just like his moral psychology, his congruence argument in *Theory* and his overlapping consensus argument in *Political Liberalism*, Rawls's argument from the absence of special psychologies is also grounded in Rousseau's concept of amour-propre. The upshot of this chapter is that there is a Rousseauvian interpretation of Rawls's stability argument. On this interpretation, some of the objections against Rawls can be answered. Thus, to fully and adequately understand Rawls, we must read Rawls in a Rousseauvian way.

CHAPTER 5 THE SIGNIFICANCE OF ROUSSEAU'S CONCEPT OF AMOUR-PROPRE IN RAWLS'S LAW OF PEOPLES

In the previous two chapters, I have argued that there is a Rousseauvian interpretation of Rawls's justice as fairness and stability argument. There is a Rousseauvian interpretation of Rawls's justice as fairness because first, one of the main reasons why parties in the original position would choose justice as fairness over other principles of justice is that justice as fairness meets the demand of amour-propre; and second, Rawls's argument from self-respect for the two principles of justice parallels, and is grounded in, Rousseau's thoughts on amour-propre. There is a Rousseauvian interpretation of Rawls's stability argument for three reasons. First, Rawls's moral psychology is an illustration of Rousseau's thesis. Second, Rawls's congruence argument in *Theory* and his overlapping consensus argument in *Political Liberalism* is grounded in Rousseau's conception of persons. And Third, Rawls's argument from the absence of special psychologies is grounded in Rousseau's concept of amour-propre. These two arguments are not separate from each other. Or we may say that Rousseau's concept of amour-propre bridges Rawls's justice as fairness and stability. One of the main reasons why citizens in the well-ordered society of justice as fairness willingly comply with the two principles of justice and stabilize the society is that justice as fairness meets the demand of amour-propre.

If plausible, these two arguments establish that there is a Rousseauvian interpretation of Rawls's theory of domestic justice. But it remains to be seen whether there is a Rousseauvian interpretation of Rawls's theory of international justice. Rawls says that his theory of international justice is consistent with his domestic theory of justice, as he says: "I consider how the content of the law of peoples might be developed out of a liberal idea of justice similar to, but more general than, the idea I called justice as fairness" (LP 3). But it has been debated whether

Rawls's theory of international justice is consistent with his theory of domestic justice.[86] If there is a Rousseauvian interpretation of Rawls's theory of international justice, and if my interpretation in the previous chapters is plausible, Rawls's theory of international justice is consistent with Rawls's theory of domestic justice. Furthermore, Rawls's *The Law of Peoples* is the completion of *Political Liberalism* in the sense that for a well-ordered society to be stable it must also be able to return to its normal state in the face of external threats. If there is a Rousseauvian interpretation of Rawls's theory of international justice, according to which peoples have the fundamental interests of amour de soi and amour-propre which are respected and promoted by the principles of the law of peoples accepted by all liberal and decent peoples, this serves as the completion of the Rousseauvian interpretation of Rawls's stability argument.

In this chapter, I will first argue that three themes of Rawls's law of peoples, that is, the two international original positions, the idea of a realistic utopia, and the distinction between peoples and states, parallel, and are grounded in, Rousseau's concept of amour-propre. Second, I will introduce Rawls's argument that the law of peoples meets the demand of amour-propre, his stability argument in the law of peoples, and his arguments against political realism and distributive cosmopolitanism. I argue that these arguments parallel, and are grounded, in Rousseau's concept of amour-propre. And finally, I will consider two objections against Rawls's argument. The first objection says that Rawls's law of peoples neglects women's need for recognition in decent peoples. The second objection says that Rawls's law of peoples fails to provide the social bases of self-respect for other societies such as benevolent absolutism, burdened societies and outlaw states. I argue that there are Rousseauvian responses to these

[86] For example, Kuper (2000), Beitz (2000), Wenar (2001), Moellendorf (2002), Nussbaum (2002), Buchanan (2004), Pogge (2004) argue that Rawls's theory of international justice is inconsistent with his theory of domestic justice. Reidy (2004; 2007), Maffetone (2011), Kim (2015) argue that it is consistent.

objections. In response to the first objection, I argue that Rawls's treatment of decent peoples meets the demand of amour-propre for all peoples and it meets the demand of amour-propre at least in a minimal sense for all natural persons. In response to the second objection, I argue that Rawls's treatment of outlaw states and other societies meets the demand of amour-propre for all persons and that outlaw states and other societies are not moral persons that have a positive sense of amour-propre. I conclude that there is a Rousseauvian interpretation of Rawls's law of peoples. In particular, Rousseau's concept of amour-propre plays a significant role in Rawls's law of peoples.

1. Amour-propre and Rawls's Law of Peoples

There are three main themes in Rawls's law of peoples. The first is to "consider how the content of the Law of Peoples might be developed out of a liberal idea of justice similar to, but more general than, the idea I called justice as fairness in *A Theory of Justice* (1971)" (LP 3). The second is to "discuss how such a Law of Peoples fulfills certain conditions, which justify calling the Society of Peoples a realistic utopia (see §1)" (LP 4). The third is to "explain why I have used the term 'peoples' and not 'states'" (LP 4). I argue that these themes can be traced back to Rousseau, especially his concept of amour-propre.

First, in the two international original positions, the liberal conception of justice is extended to the law of peoples.[87] Rawls tries to defend the social contract tradition of Rousseau, Locke and Kant in both his domestic and international theories of justice. In the domestic case, he argues that all parties who represent free and equal, rational and reasonable beings, in the original

[87] Rawls says: "At the next level, the idea of the original position is used again, but this time to extend a liberal conception to the Law of Peoples" (LP 32).

position under the veil of ignorance would agree on the two principles of justice. Parallel to his original position argument in the domestic case, Rawls first argues that in the international case, parties that represent free and equal, rational and reasonable, liberal peoples in the first original position under the veil of ignorance would choose the eight principles of the law of peoples. Under the veil of ignorance, parties in the first international original position do not know the size of their territory, the population, the extent of their natural resources, the level of their economic growth and so on. They only know that they represent free and equal, rational and reasonable liberal peoples and that the peoples they represent have two fundamental interests: one is political independence, the security of its territory and the well-being of its people; another is amour-propre. Thus stipulated, Rawls argues that the parties would agree on the following principles:

(1) Peoples are free and independent, and their freedom and independence are to be respected by other peoples.

(2) Peoples are to observe treaties and undertakings.

(3) Peoples are equal and are parties to the agreements that bind them.

(4) Peoples are to observe a duty of non-intervention.

(5) Peoples have the right of self-defense but no right to instigate war for reasons other than self-defense.

(6) Peoples are to honor human rights.

(7) Peoples are to observe certain specified restrictions in the conduct of war.

(8) Peoples have a duty to assist other peoples living under unfavorable conditions that prevent their having a just or decent political and social regime.

Rawls then argues that it is unreasonable to suppose that all nonliberal nondemocratic societies are not well-ordered. He argues that the law of peoples can be extended to decent peoples because he conjectures that in the second international original position, decent peoples who are rational and responsible would also agree on the eight principles of justice. Furthermore, he thinks that it would hurt their self-respect if they were not included in the society of peoples.

Thus, the first theme of Rawls's law of peoples is Rousseauvian for three reasons. First, just like in the domestic case, his international original position argument follows the social contract tradition of Rousseau, Locke and Kant. Second, one of the fundamental interests of the peoples in the society of peoples is their proper amour-propre. As I will explain in detail later, meeting the demand of amour-propre is one of the main reasons why parties in the international original position would choose the law of peoples. Third, one of the main reasons why Rawls includes decent peoples in the society of peoples is that denying them membership would hurt their amour-propre.

Second, Rawls argues that his Society of Peoples, just like a liberal society, is a realistic utopia. There are several conditions for a liberal society to be realistic. The first condition is that "it must rely on the actual laws of nature and achieve the kind of stability those laws allow, that is, stability for the right reasons" (LP 12). Rawls acknowledges that here he draws on Rousseau's thought in *The Social Contract*, that is, "taking men as they are and laws as they might be" (LP 13). The second condition is that "it is workable and may be applied to ongoing cooperative political arrangements and relations between peoples" (LP 17). It is workable because citizens share the desire for the social primary goods which include the social bases of self-respect. We have seen that Rawls's idea of self-respect as perhaps the most important primary good parallels Rousseau's story of the unhappy fugitive who lacks self-esteem. There are several conditions for

a conception of justice to be utopian. First, it must specify a reasonable and just society in which the social primary goods are equally or fairly distributed. As we have seen, one of the main reasons why Rawls's justice as fairness is just is that the equal or fair distribution of social primary goods secures the social bases of self-respect. This also explains why Rawls's justice as fairness is Rousseauvian. The second reason for a conception of justice to be utopian is that it satisfies the criterion of reciprocity. We have seen that the principle of reciprocity is originated from the reciprocal feature of Rousseau's concept of amour-propre. The third condition for a realistic utopia is that it must have the essential elements of political liberalism. As we have seen, for both Rousseau and Rawls, a conception of justice by a comprehensive doctrine is not the correct path toward a society of mutual-respect and self-esteem. The fourth condition is that liberal democracies must have political and social institutions that lead their citizens to a sufficient sense of justice. This brings stability to these societies. We have seen that Rawls's moral psychology is an illustration of Rousseau's thesis and that Rawls's stability argument is grounded in Rousseau's concept of amour-propre. The fifth condition is that the stability of liberal democracies is rooted in a political conception of justice affirmed by an overlapping consensus of all reasonable comprehensive doctrines. We have seen that the overlapping consensus argument is grounded in Rousseau's conception of political freedom. The sixth condition is that the political conception of justice must "show the reasonableness of toleration by public reason" (LP 16). We have also seen that Rawls's idea of public reason parallels Rousseau's idea of the general will. Since Rawls's law of peoples is similar to a liberal society in

the sense that they are both realistically utopian, the second theme of Rawls's law of peoples is also Rousseauvian.[88]

Third, Rawls argues that peoples are different from states. For Rawls, persons are subjects of claims and so persons are not limited to individual human persons (LP 75). Nations, Corporations, churches, teams are all persons that have claims on others. Peoples, in this sense, are also persons. As moral persons, peoples see themselves as free and equal, and as having fundamental interests of protecting their political autonomy, security, territory, and the well-being of its people. They also see themselves as having the fundamental interest of self-respect, or amour-propre. Since peoples have a secure sense of self-respect, they are prepared to respect other peoples and to recognize their equality while states do not. Rawls says: "What distinguishes peoples from states — and this is crucial — is that just peoples are fully prepared to grant the very same proper respect and recognition to other peoples as equals" (LP 35). And so the crucial difference between states and peoples is that the former is merely rational while the latter is also reasonable. The proper sense of amour-propre inclines peoples to act reasonably, that is, "to grant the very same proper respect and recognition to other peoples as equals." To specify, outlaw states are not reasonable because they have an inflamed amour-propre that inclines them to dominate other societies; burdened societies are not reasonable because they do not have the resources to be well-ordered, or to be reasonable; benevolent absolutisms are not reasonable because they are systems of coordination rather than systems of cooperation which are ready to grant proper respect and recognition to other peoples. Thus, the reciprocal feature of

[88] In the next section, I will present several arguments such as Rawls's moral psychology, his stability argument, and his overlapping consensus argument in the law of peoples. I argue that these arguments parallel, and are grounded in, Rousseau's concept of amour-propre, and so the second theme parallels, and is grounded in, Rousseau's concept of amour-propre is reaffirmed.

Rousseau's concept of amour-propre explains the crucial distinction between peoples and states. And so the third theme of Rawls's law of peoples is also Rousseauvian.

To sum up, the three themes of Rawls's law of peoples are Rousseauvian. The first theme of Rawls's law of peoples parallels, and is grounded, in Rousseau's concept of amour-propre because as we will later see, meeting the demand of amour-propre is one of the main reasons why Rawls's law of peoples is chosen by the parties in the international original position. The second theme of Rawls's law of peoples parallels, and is grounded in, Rousseau's concept of amour-propre because Rawls's idea of a realistic utopia parallels, and is grounded in, Rousseau's concept of amour-propre. The third theme of Rawls's law of peoples parallels, and is grounded in, Rousseau's concept of amour-propre because the crucial distinction between peoples and states is grounded in the reciprocal feature of amour-propre.

2. Rawls's Law of Peoples Meets the Demand of Amour-propre

To specify why the first theme of Rawls's law of peoples parallels, and is grounded in, Rousseau's concept of amour-propre, I argue that one of the main reasons why the principles of the law of peoples are chosen by the parties in the international original position is that Rawls's law of peoples meets the demand of amour-propre. I first explain why Rawls's law of peoples meets the demand of amour-propre. Then, I argue that Rawls's stability argument in the law of peoples parallels, and is grounded in, Rousseau's concept of amour-propre. And finally, I consider Rawls's objections against political realism and the distributive cosmopolitanism. I argue that political realism does not really have a moral psychology and that distributive cosmopolitanism is not necessary for (and fails to meet) the demand of amour-propre.

2.1 Why Does Rawls's Law of Peoples Meet the Demand of Amour-propre?

We have seen that in the domestic case, one of the main reasons why parties who are free and equal rational beings would choose Rawls's justice as fairness over other principles of justice is that Rawls's justice as fairness meets the demand of amour-propre. And this also brings about the stability of the well-ordered society of justice as fairness. Similar to the domestic case, one of the main reasons why parties in the international original position would choose Rawls's law of peoples is that the law of peoples meets the demand of amour-propre. But different from the domestic case, there are two kinds of persons that Rawls considers in the law of peoples. The first kind of person is peoples. The second kind is natural persons. I argue that Rawls's law of peoples meets the demand of amour-propre in the international case in two ways. First, I argue that Rawls's law of peoples takes into account the amour-propre of reasonable peoples, whether liberal and democratic or decent. Second, I argue that Rawls's law of peoples takes into account the amour-propre of natural human persons. For one thing, it requires a condition of admission into the society of peoples that polities be organized so that basic human rights are secured and thus that their members are at least recognized and respected as rational and responsible. While this recognition and respect are insufficient for those who have developed into Stage 3 of moral development, a polity that respect these human rights is no obstacle for its members to reform it to maintain a proper sense of amour-propre. For another thing, the international community is to respect and honor the equal human rights for all natural persons which include those who live in outlaw states, burdened societies and benevolent absolutisms, and so this also provides the social bases of self-respect for all natural persons in a minimal sense. Similar to the domestic case, meeting the demand of amour-propre brings about the stability of the Society of Peoples.

To specify, there are two main reasons why Rawls's law of peoples meets the demand of amour-propre. First, the freedom, independence and equality of the peoples, either liberal and democratic or decent, are respected by other peoples in the law of peoples. We have seen that amour-propre is the desire for equality and freedom. The social bases of amour-propre for citizens are secured by the public affirmation of the two principles of justice. The social bases of amour-propre for peoples are (partly) secured by the public recognition of the law of peoples. There are several principles of the law of peoples that manifestly secure the freedom, equality and independence of the members of the society of peoples: "1. Peoples are free and independent, and their freedom and independence are to be respected by other peoples" (LP 37); and "3. Peoples are equal and are parties to the agreements that bind them" (LP 37); and "4. Peoples are to observe a duty of non-intervention" (LP 37); and "5. Peoples have the right of self-defense but no right to instigate war for reasons other than self-defense" (LP 37). These rights and liberties are publicly affirmed through the law of peoples, and so Rawls's law of peoples provides the social bases of self-respect for both liberal democratic peoples and decent peoples. Thus Rawls's law of peoples meets the demand of amour-propre for both liberal democratic and decent peoples.

Second, Rawls's law of peoples also meets the demand of amour-propre for natural persons in a minimal sense.[89] For one thing, it takes into account the amour-propre of natural human persons by requiring a condition of admission into the society of peoples that polities be

[89] I argue that Rawls's list of human rights meets the demand of amour-propre for all natural persons at least in the minimal sense. What does it mean that it meets the demand of amour-propre in the minimal sense? Recall Rousseau's story of the unhappy fugitive who lacks self-esteem (in Chapter 2). This young expatriate changed his religion for bread, and was taught new dogma and evils. As he wanted to flee, he was locked up. The ecclesiastic helped him escape and gave him a fortune but he soon lost everything. He was full of self-contempt by that time. To have self-respect in a minimal sense, therefore, means that a person should have the right to life (which includes the right to subsistence), the right to liberty (which include freedom of religion), the right to property, and the right to formal equality (which includes the freedom from arbitrary treatment) and so on.

organized so that basic human rights are secured and thus that their members are at least recognized and respected as rational and responsible. For Rawls, a special class of urgent rights[90] is to be respected by liberal and decent peoples in Rawls's law of peoples.[91] The sixth principle of the law of peoples manifests this aspect of the law of peoples: "6. Peoples are to honor human rights" (LP 37). These human rights include the rights to life, liberty, property and formal equality (LP 65).[92] Since a polity has to respect these human rights to be a member of the Society of Peoples, Rawls's law of peoples provides the social bases of self-respect for the members of both liberal democratic and decent peoples. Even though not everyone is granted equal political rights in decent peoples, they all have the right to liberty which includes "freedom from slavery, serfdom, and forced occupation, and to a sufficient measure of liberty of conscience to ensure freedom of religion and thought" and so on (LP 65). The public recognition of these human rights provides the social bases of self-respect for the members of decent peoples because they are thus respected and recognized as rational and responsible. While these human rights are insufficient for them to have a secure sense of amour-propre, a polity that respects these human

---

[90] Rawls says that "Human rights in the Law of Peoples, by contrast, express a special class of urgent rights, such as freedom from slavery and serfdom, liberty (but not equal liberty) of conscience, and security of ethnic groups from mass murder and genocide" (LP 79). I think that the reason why these rights are "urgent rights" for Rawls is that the violation of these rights is extremely corrosive of self-respect. And so Rawls's list of human rights is essential for the demand of amour-propre to be met.

[91] This brings about the distinction between moral cosmopolitanism and distributive cosmopolitanism. The cosmopolitanism that Rawls objects to is distributive cosmopolitanism. David Held and Pietro Maffettone (2016, p.7) points out that Rawls and his opponents are all moral cosmopolitans because they agree that every person should be treated with equal respect. But Rawls and his opponents disagree about what it means to treat every person with equal respect. I argue that the law of peoples treats every person with equal respect in two ways: first, the law of peoples treats all peoples (who are also moral persons) with equal respect; and second, it treats every human being with equal respect by honoring human rights.

[92] David Reidy (2006, p.170) points out that Rawls's list of basic human rights "is less minimalist than many critics have allowed." He argues that the rights specified in Articles 3-18 of the UDHR (Universal Declaration of Human Rights) are "the full and most fundamental sense of the term." These rights include "the central elements of due process and the rule of law (Articles 6-12 and 17), the right to refuse nonconsensual marriage (Article 16), a right against cruel, inhuman or degrading punishment and against torture (Article 5), the right to seek asylum (Article 14), the right to a national identity (Article 15), and the right to freedom of movement (Article 13)." Other rights that are entailed by these rights can also be included in the list of human rights.

rights is no obstacle for its members to reform it to have a secure sense of amour-propre. For another thing, and more importantly, Rawls thinks that the natural persons in outlaw states, burdened societies and benevolent absolutisms are to be respected and recognized as having equal human rights as others. Even though burdened societies and outlaw states do not respect the human rights of their members, these human rights are still recognized and respected by liberal and decent peoples (or the international community). And this also serves as incentives for the members of these societies to reform them to maintain a secure sense of amour-propre. The second, seventh and eighth principles of the law of peoples manifest this aspect of the law of peoples: "2. Peoples are to observe treaties and undertakings" (LP 37); and "7. Peoples are to observe certain specified restrictions in the conduct of war. 8. Peoples have a duty to assist other peoples living under unfavorable conditions that prevent their having a just or decent political social regime" (LP 37). The second principle regulates international institutions, the standards of fair trade and so on.[93] The seventh principle restricts the violation of human rights to extremely severe conditions under which the existence of liberal democracies is at stake. The eighth principle specifies the duty of assistance that helps burdened societies to be well-ordered. These principles show that Rawls takes the amour-propre of the natural persons in non-well-ordered societies into account directly. Thus, Rawls's law of peoples meets the demand of amour-propre for all natural persons in a minimal sense.

Therefore, Rawls's law of peoples meets the demand of amour-propre for two reasons. The first reason is that the law of peoples ensures the social bases of self-respect for both liberal

[93] For a discussion of what it means to have fair trade, see Aaron James (2016). James (2016) argues for a world-historical approach to explain fair trade instead of an interpersonal approach. It seems to me, however, that Rawls's sixth principle of the law of peoples comes mainly from the consideration of treating every human being with equal respect. And so Rawls's explanation of fair trade is more likely to fall into an interpersonal approach range.

democratic and decent peoples in the society of peoples. The second reason is that the law of peoples honors the most urgent human rights that are essential for a minimal sense of amour-propre for all natural human persons. This also give incentives for the members of non-liberal non-democratic societies to reform them to maintain a secure sense of amour-propre.

2.2 Amour-propre and Stability in the Law of Peoples

In the domestic case, Rawls argues that the stability of the well-ordered society of justice as fairness is achieved for two reasons. First, citizens acquire the sense of justice in Stage 3 of moral development. Second, there is an overlapping consensus from all reasonable comprehensive views. Since the fundamental interests of citizens, amour de soi and amour-propre, are met by the principles of justice, citizens willingly comply with the principles of justice. I have argued that Rawls's moral psychology and the overlapping consensus argument in the domestic case parallel, and are grounded in, Rousseau's concept of amour-propre. Similarly, in the international case, Rawls argues that the stability of the society of peoples is achieved for two reasons. First, peoples acquire the sense of justice in Stage 3 of moral development. Second, since the fundamental interests of peoples, amour de soi and amour-propre, are met by the law of peoples, peoples willingly comply with the principles of the law of peoples. I argue that Rawls's moral psychology and the overlapping consensus argument in the international case parallel, and are grounded, in Rousseau's concept of amour-propre.

2.2.1 Amour-propre and Rawls's Moral Psychology in the Law of Peoples

Rawls argues that the stability of the society of peoples is not the stability as a balance of power, but the stability for the right reason. He argues that a liberal society is likely to be stable

for the right reason because citizens acquire the sense of justice that inclines them to accept the principles of justice. Similar to the domestic case, he argues that the society of peoples is likely to be stable for the right reason:

> once the second original position argument is complete and includes *the account of moral learning*, we conjecture, first, that the Law of Peoples the parties would adopt is the law that we — you and I, here and now — would accept as fair in specifying the basic terms of cooperation among peoples. (LP 45)[94]

We have seen that Rawls's moral psychology is an illustration of Rousseau's thesis that "the sense of justice is no mere moral conception formed by the understanding alone, but a true sentiment of the heart enlightened by reason, the natural outcome of our primitive affections" (CP 96). In the domestic case, Rawls argues that citizens acquire the sense of justice in Stage 3 of moral development if they see that the just institution that they live in is beneficial to them and those around them, given that they have acquired love and friendship in the first two stages. He argues that the sense of justice inclines the citizens to accept and act on and from the principles of justice. Similarly, here Rawls argues that peoples as moral persons acquire the sense of justice in stage 3 of moral development, if they see that the law of peoples is beneficiary to them and those around them, given that they have acquired love and friendship in the first two stages.[95] He

---

[94] Italics added.

[95] What does it mean to say that peoples have the sense of justice at Stage 3 of moral development, if they see that the law of peoples is beneficiary to them and those around them, given that they have acquired love and friendship in the first two stages? For Rawls, peoples, like persons, have achievements and accomplishments in history and they base their self-respect on them. Thus, to say that peoples, like persons, experience three stages of moral development, is to say that their domestic political and social institutions develop from benevolent absolutism (and other societies) to decent and liberal peoples. As David Reidy (forthcoming, p.20) points out "Rawls's account of the benevolent absolutism, decent constitutional republic and liberal democracy tracks his account of moral development from *Theory*."

argues that the sense of justice inclines peoples to accept and act on and from the principles of the law of peoples.

We have seen that amour-propre plays a significant role in Rawls's moral psychology in the domestic case. Since peoples, like persons, also acquire the sense of justice in stage 3 of moral development that inclines them to accept the law of peoples, amour-propre also plays a significant role in Rawls's moral psychology. Thus, there is a Rousseauvian interpretation of Rawls's moral psychology in the international case.

2.2.2 Amour-propre and Rawls's Overlapping Consensus Argument in the Law of Peoples

Another reason why peoples are willing to honor the principles of the law of peoples is that their fundamental interests are respected and promoted by the law of peoples. As Rawls writes:

> When the Law of Peoples is honored by peoples over a certain period of time, with the evident intention to comply, and these intentions are mutually recognized, these peoples tend to develop mutual trust and confidence in one another. Moreover, peoples see those norms as advantageous for themselves and for those they care for, and therefore as time goes on they tend to accept that law as an ideal of conduct. (LP 44)

Rawls says that peoples are rational and reasonable and that they have the two fundamental interests. One of the fundamental interests is "to protect their political independence and their free culture with its civil liberties, to guarantee their security, territory, and the well-being of their citizens" (LP 34). In Rousseau's terms, peoples, like persons, are partly moved by

amour de soi. Another fundamental interest is amour-propre, or the desire to receive "from other peoples a proper respect and recognition of their equality" (LP 35). Thus, there is an overlapping consensus among all liberal and decent peoples, that is, to honor the law of peoples that protect and promote their fundamental interests, amour de soi and amour-propre. In other words, given that peoples have acquired the sense of justice in the third stage of moral development, once they see that the law of peoples respects and promotes their fundamental interests, they willingly comply with the principles of the law of peoples.

Therefore, Rousseau's concept of amour-propre also plays a significant role in Rawls's overlapping consensus argument in the international case. And so there is a Rousseauvian interpretation of Rawls's overlapping consensus argument in the law of peoples.

2.3 Amour-propre and Rawls's Arguments against Political Realism and Distributive Cosmopolitanism

We have seen that in the domestic case Rawls's moral psychology is a realistic moral psychology. This explains why his moral psychology is more understandable or more realistic than the moral psychologies of intuitionism, perfectionism and utilitarianism. In the international case, there are no rival conceptions of international justice for the parties in the original position to choose from. Rawls's arguments against political realism and distributive cosmopolitanism explains why they cannot be presented to the parties in the international original position to choose from. I argue that amour-propre plays a significant role in Rawls's arguments against political realism and distributive cosmopolitanism.

2.3.1 Amour-propre and Rawls's Argument against Political Realism

Political realism holds that the participators in international relations only aim to advance their rational self-interest and that the stability can only be achieved by a balance of forces. In response to political realism, Rawls writes: "Part of the answer to political realism is that this reasonable sense of proper respect is not unrealistic, but is itself the outcome of democratic domestic institutions" (LP 35). Peoples are moral persons. The moral nature of peoples includes a proper sense of amour-propre that rests on their achievements and accomplishments in their history, on the freedom and integrity of their citizens, and on the justice and decency of their domestic institutions (LP 47). Since amour-propre has a reciprocal feature, peoples who have amour-propre will have the reciprocity of disposition that inclines them to accept and act on and from the principles of the law of peoples. And so political realism is mistaken to suppose that all nations are only rational and that stability can only be achieved by a balance of forces.

Therefore, amour-propre plays a significant role in Rawls's argument against political realism. Political realism fails to recognize the reciprocal feature of amour-propre and so it does not really provide a *moral* psychology.

2.3.2 Amour-propre and Rawls's Arguments against Distributive Cosmopolitanism

The distributive cosmopolitan view starts from the premise that all natural persons are rational and reasonable and that they have the capacity for a conception of the good and the capacity for the sense of justice. According to distributive cosmopolitanism, in the international original position argument, all parties would choose the first principle of justice that grants equal basic liberties and rights for all natural persons. Rawls argues that the distributive cosmopolitan view is unrealistic, as he says:

> On this account, the foreign policy of a liberal people — which it is our concern to elaborate — will be to act gradually to shape all not yet liberal societies in a liberal direction, until eventually (in the ideal case) all societies are liberal. But this foreign policy simply assumes that only a liberal democratic society can be acceptable. (LP 82)

Rawls argues that it is simply unrealistic to assume that liberal democracies are the only acceptable societies. This is for the reason that Rawls conjectures that decent peoples are ready and willing to interact with liberal democratic peoples on terms that liberal democratic peoples recognize as reasonable. He thus argues that nonliberal or nondemocratic decent peoples are owed due respect by liberal peoples. He proposes an argument from amour-propre in favor of this conjecture, that is, if decent peoples are not respected by liberal peoples, "This lack of respect may wound the self-respect of decent nonliberal peoples as peoples, as well as their individual members, and may lead to great bitterness and resentment" (LP 61). Thus, liberal peoples who have a secure sense of amour-propre are ready to grant due respect to decent peoples, even though they believe that they, as liberal democracies, are superior to other forms of societies (LP 62).[96]

The distributive cosmopolitanism view can take two forms. It can focus either on the distribution across the body politics in the Society of Peoples or across natural persons. It argues for the global distributive principles of justice such as the resource redistribution principle and the global egalitarian principle. The resource redistribution principle says that the redistribution of resources should make sure that resource-poor countries have the same opportunity to meet the basic needs of its people as resource-rich countries. Rawls argues that this principle

[96] Rawls notes that this proper self-respect does not depend on their being superior or inferior to other societies (LP 48).

undermines the political autonomy of the peoples in the society of peoples. The global egalitarian principle proposes a General Resource Dividend (GRD) for each society to pay an international fund to support the global poor. Rawls argues that the problem with this principle is that it does not have a target or cutoff (LP 119).

What underlies Rawls's argument against distributive cosmopolitanism is the demand of amour-propre. First, it is not necessary to have the global distributive principles to meet the demand of amour-propre in the international case. Like the domestic case, the demand of amour-propre is met if equal basic liberties and rights are publicly recognized and respected in the international case. And so Rawls writes that the equality, freedom and independence of peoples must be respected by other peoples and that basic human rights must be honored by all peoples. In these ways, the demand of amour-propre is met for all peoples and all natural persons. But unlike the domestic case in which the fair equality of opportunity principle and the difference principle are proposed partly for everyone to have equal basic rights and liberties substantively, in the international case, Rawls does not argue for such principles. This is partly for the reason that once a society becomes well-ordered, it will respect and honor the equality, freedom and independence of other well-ordered societies, and the basic human rights for every natural person. Its amour-propre does not rest on its having more resources or wealth than other countries, but on its domestic public culture, its accomplishment and achievement in its history, and its freedom and integrity as a people. Thus, Rawls argues for the principle of assistance, which has the target or cut off, rather than the global distributive principles, which does not have the target or cut off, to assist burdened societies to be well-ordered so that they can have sufficient self-respect that inclines them to respect other peoples in the international case.

Second, the global distributive principles emphasize the importance of economic equality across the body politics or natural persons. This may take two forms: the first one is to argue for the global egalitarian principle without the priority of equal liberty; the second one is to argue for the global egalitarian principle with the priority of equal liberty. I argue that either way, it leads to an inflamed amour-propre across the body politics or natural persons. First, the global egalitarian principle without the priority of equal liberty undermines the self-respect of peoples and natural persons. in the domestic case, Rawls argues that the legitimate social bases of self-respect are not socio-economic status but equal citizenship. He argues that grounding a person's self-respect on the socio-economic status necessarily undermines his self-esteem. This is for the reason that if his socio-economic status is lower than others, grounding his self-respect on the socio-economic status will make him feel that he is inferior to others. Now since the global egalitarian principle (such as the global difference principle) still allows large economic inequalities across body politics or natural persons, if peoples or natural persons ground their self-respect merely on economic equality, they will necessarily have an inflamed amour-propre that inclines them to dominate other peoples or natural persons. But if economic equality is strictly required by the global distributive principles, some of the basic liberties or rights might be undermined for the sake of economic equality. Sacrificing basic rights or liberties for the sake of economic equality undermines the self-respect of the peoples. Second, the global egalitarian principle with the priority of equal liberty also undermines the self-respect of peoples and natural persons. If distributive cosmopolitanism argues that the global difference principle is required to grant everyone equal basic rights and liberties substantively, Rawls would say that such a principle is neither necessary nor desirable. We have seen the reason why such a principle is not necessary — once a society is well-ordered, it will have a sufficient sense of amour-propre that

rests on its moving to decency or liberal democracy. Such a principle is also not desirable. In the domestic case, Rawls's two principles of justice apply to the basic structure of the well-ordered society of justice as fairness; but in the international case, there is no international basic structure for the difference principle to apply. If a global legislature is made possible in the future, Rawls predicts that such a global legislature is either impotent or despotism. If the global legislature is impotent, the global difference principle is not applicable; but if it is despotism, the political autonomy of the peoples and their members is undermined, and so is their self-respect. Thus, the global distributive principles cannot meet the demand of amour-propre.

To sum up, Rawls's law of peoples meets the demand of amour-propre because first, the freedom, independence and equality of peoples are respected by both liberal democratic and decent peoples; and second, the most urgent human rights of every human being are respected by all peoples in the society of peoples, and this also gives incentives for the members of non-liberal and non-democratic societies to reform them to maintain a secure sense of amour-propre. Furthermore, since Rawls's law of peoples meets the demand of amour-propre, peoples (and individuals) are motivated to stabilize the society of peoples. Rawls's stability argument in the law of peoples also relies on Rousseau's account of moral education, especially, the reciprocal feature of amour-propre. Political realism and distributive cosmopolitanism are two rival conceptions of international justice that Rawls considers in the law of peoples. In response to political realism, Rawls argues that the law of peoples relies on a realistic moral psychology. In response to the distributive cosmopolitanism, Rawls argues that it is unnecessary to have the global distributive principles to meet the demand of amour-propre and that the global distributive principles fail to meet the demand of amour-propre. Therefore, there is a Rousseauvian

interpretation of Rawls's law of peoples. In particular, Rousseau's concept of amour-propre plays a significant role in Rawls's law of peoples.

3. Objections and Responses

Two objections are raised against Rawls's argument that the law of peoples meets the demand of amour-propre. First, Rawls's law of peoples presumably provides the social bases of self-respect for the peoples in the society of peoples, but it does not provide the social bases of self-respect for some members in the society of peoples. In particular, in a decent hierarchical society such as Kazanistan, women are denied some of the important political rights. For example, they may not have the right to vote. Since they are denied equal citizenship in the decent society, they will not have a secure sense of self-respect.[97] Second, Rawls's law of peoples presumably provides the social bases of self-respect for the peoples in the Society of Peoples, but it does not provide the social bases of self-respect for benevolent absolutism, burdened states and outlaw states which also have the need for equal membership in international relations.[98] Denying the equal membership of these states would wound their self-respect and lead to bitterness and resentment. This may bring instability to the Society of Peoples.

I will argue that there are Rousseauvian responses to both questions. To the first question, I argue that Rawls's treatment of decent peoples meets the demand of amour-propre for all peoples and that it meets the demands of amour-propre at least minimally for all persons. To the second question, I argue that denying them membership in the law of peoples meets the demand of amour-propre for all persons. I also argue that one of the main reasons Rawls excludes burdened

[97] This objection is raised by Nussbaum (2002).
[98] This objection is raised by Beitz (2000).

societies, benevolent absolutisms and outlaw states in the law of peoples is that they are not moral persons who have a positive form of amour-propre.

3.1 The First Objection

In response to the first objection, there are several existing interpretations as to why Rawls neglects women's need for recognition in the decent peoples. First, the practice-dependent interpretation endorsed by, among others, Aaron James (2005), Leif Wenar (2006), David Reidy (2004) and Pietro Maffettone (2011), argues that our current social practices such as international laws and institutions are only participated in by nations and that they represent nations instead of individuals as free and equal. While this interpretation explains why Rawls's law of peoples must meet the demand of amour-propre for all peoples, it fails to explain why it must also meet the demand of amour-propre for all persons. Gillian Brock (2010, p.95) argues that the interdependence of the global political culture indicates that individuals should also be the focus of global justice. She argues that "Every member state of the United Nations has signed one or more of the six major human rights treaties, with more than 80 percent having ratified four or more of these six core documents" (ibid.). What Brock says is not an objection against Rawls's law of peoples as she proposes to be, but an objection against practice-dependent interpretation of Rawls's law of peoples. As we have seen that Rawls's law of peoples takes the amour-propre of natural persons into account, and so it takes the interdependence of the global political culture seriously. But Brock's argument can be seen as an objection against the practice-dependent interpretation which does not recognize the fact of the interdependence of the global political culture.

Second, the stability interpretation proposed by Hyunseop Kim (2015) argues that Rawls does not simply endorse the status quo; he argues that Rawls assumes that domestic public political culture is the main cause of international security and peace, and so liberal and decent peoples do not go to war with each other. Kim seems to suggest that since Rawls's main aim in *The Law of Peoples* is to complete his stability argument in *Political Liberalism,* he needs not to address the problem of low self-respect for individuals. But the stability interpretation does not really explain why Rawls needs not to address this problem from the consideration of stability. Allen Buchanan (2000) argues that if some individuals lack self-respect in decent peoples, it will bring about the intra-state conflicts that undermine the stability of the Society of Peoples. Buchanan's objection can be seen as an objection against Kim's stability interpretation of Rawls, but it is not an objection against Rawls's law of peoples as he proposes to be. As we have seen that Rawls's law of peoples takes into account of the amour-propre of the members in decent peoples. They are respected and recognized as rational and responsible and their societies are no obstacles for them to reform them to maintain a secure sense of amour-propre. And so since the demand of amour-propre of the members of the decent peoples is met to some extent, even though the decent societies are not as stable as liberal democracies (at least not in the sense of constitutional stability), these societies are stable enough not to undermine the stability of the Society of Peoples.

My interpretation, unlike the practice-dependent interpretation and the stability interpretation, argues that Rawls's law of peoples meets the demand of amour-propre for all peoples, and that it meets the demand of amour-propre at least minimally for all persons. Rawls realizes that the satisfaction of the demand of amour-propre for all peoples may be in conflict with the satisfaction of the demand of amour-propre for all persons. And so even though his law

of peoples tries to meet both principles, when they conflict with each other, he tries to give strong reasons to give priority to one of the principles. For example, in arguing that liberal peoples must respect decent peoples, he argues that there are no strong reasons not to give priority to the principle that the law of peoples must meet the demand of amour-propre for all peoples.[99] And in arguing that outlaw states must not be tolerated by liberal and decent peoples, he gives strong reasons for giving priority to the principle that the law of peoples must meet the demand of amour-propre for all persons.[100]

Rawls argues that liberal peoples must respect decent peoples partly because there are no strong reasons not to respect them. Rawls argues that liberal peoples cannot simply assume that there are no nonliberal and nondemocratic peoples that deserve moral recognition and respect. Rawls says:

> Just as citizen in a liberal society must respect other persons' comprehensive religious, philosophical and moral doctrines provided they are pursued in accordance with a reasonable political conception of justice, so a liberal society must respect other societies organized by comprehensive doctrines, provided their

---

[99] Rawls writes: "Denying respect to other peoples and their members requires strong reasons to be justified. Liberal peoples cannot say that decent peoples deny human rights, since (as we shall see in §§8-9 where the notion of decency is developed) such peoples recognize and protect these rights; nor can liberal peoples say that decent peoples deny their members the right to be consulted or a substantial political role in making decisions, since the basic structure of these societies will be seen to include a decent consultation hierarchy or its equivalent. Finally, decent peoples allow a right of dissent, and government and judicial officials are required to give a respectful reply, one that addresses the merits of the question according to the rule of law as interpreted by the judiciary. Dissenters may not be dismissed as simply incompetent or lacking in understanding. In this and other ways, the common good conception of justice held by decent peoples may gradually change overtime, prodded by the dissents of members of these peoples" (LP 61).

[100] Rawls writes: "It may be asked by what right well-ordered liberal and decent peoples are justified in interfering with an outlaw state on the grounds that this state has violated human rights…. Liberal and decent peoples have extremely good reasons for their attitude. Outlaw states are aggressive and dangerous; all peoples are safer and more secure if such state change, or are forced to change their ways" (LP 81).

> political and social institutions meet certain conditions that lead the society to adhere to a reasonable law of peoples. (LP 42-3)

It is important to note that decent peoples are to be tolerated only if "their political and social institutions meet certain conditions that lead the society to adhere to a reasonable law of peoples." Liberal peoples should tolerate decent peoples because they are all members of the Society of Peoples who accept the principles of the law of peoples. Rawls assumes that two criteria need to be met by decent peoples. The first criterion is that "the society does not have aggressive aims, and it recognizes that it must gain its legitimate ends through diplomacy and trade and other ways of peace" (LP 64). The second criterion has three parts. The first part is that "a decent hierarchical people's system of law, in accordance with its common good idea of justice (see §9), secures for all members of the people what have come to be called human rights" (LP 65). The second part is that the system of law must impose duties and obligations to all persons in the decent people. The third part is that judges and other officials must sincerely believe that the law accords with the common good of the people.

Rawls's criterion of decency is criticized as both demanding too much and demanding too little.[101] It demands too much because it excludes benevolent absolutism, burdened societies and outlaw states in the Society of Peoples. It demands too little because Rawls's list of human rights is minimal.[102] The first criticism is similar to the second objection that I will consider next. The second criticism can be discussed here. Nussbaum (2002) argues that Rawls's minimal list of human rights allows decent hierarchical societies to be members of the Society of Peoples. These

---

[101] See, for example, Tarek Hayfa (2004).
[102] For this objection, see for example, Nussbaum (2002), Alistair Macleod (2006) and Allen Buchanan (2006). For a defense of Rawls's view, see for example, Wilfried Hinsch and Markus Stepanians (2006) and David Reidy (2006).

societies do not have political democracy, the equal worth of liberty, universal suffrage, and freedom of speech. She invites us to imagine an independent state, Kerala, where women are denied the right to vote. She argues that even though these women are represented by a body in the decent consultation hierarchy, the representatives may not dare to speak out. She argues that these women are subordinated in the decent hierarchical society and that there is no reason to think that they endorse this way of life that keeps them subordinated.

In response to Nussbaum, I think that if in the independent state of Kerala women are subordinated and they feel that they are subordinated, there is no reason to include such a state in the Society of Peoples. I think Rawls would also agree that if women in a state like Kerala are subject to injustice as Nussbaum describes, there is no reason to call it "a decent hierarchical society". But there is no reason to think that women in decent peoples do not have at least a minimal sense of self-respect in so far as their human rights are respected by all peoples in the international community, and by its own societies. Furthermore, since women have some political rights that allow them to participate (unequally) in the social cooperation of the decent societies, I tend to think that they have more than a minimal sense of self-respect. In any case, if those who have developed into Stage 3 of moral development in decent peoples are not content with their unequal political rights, they may reform the political structure of the society to maintain their secure sense of amour-propre.

What are the bases of self-respect for women in decent peoples, if not equal citizenship? Rawls writes:

> A decent hierarchical society's conception of the person, as implied by the second criterion, does not require acceptance of the liberal idea of persons are citizens first and have equal basic rights as equal citizens. Rather it views persons as

> responsible and cooperating members of their respective groups. Hence, persons can recognize, understand, and act in accordance with their moral duties and obligations as members of these groups. (LP 66)

Here Rawls says that unlike persons in liberal peoples who are free and equal, reasonable and rational, persons in decent peoples are "responsible and cooperating members of their respective groups". Thus, unlike persons in liberal peoples who base their self-respect on equal citizenship, persons in decent peoples may base their self-respect on their having human rights that are publically recognized by the Society of Peoples, or on their unequal political participation in the social cooperation of the decent peoples for the common good. Even though women do not have equal citizenship in the decent consultation hierarchy, their interests are represented by a body in making final political decisions. As responsible members of the decent societies, women have their moral duties and obligations in accordance with the common good idea of justice. Assuming they are brought up to be responsible and rational, they will manifestly live up to the duties and obligations set up by the decent hierarchy. If they are dissatisfied by the decisions of the consultation hierarchy, they have the right to dissent. Their voice will be heard and they will get a respectful reply. In any case, if those who have developed into Stage 3 of moral development are not satisfied with the extent to which their demand of amour-propre is met, they can reform the political structure of the decent peoples to maintain a more robust sense of amour-propre.

3.2 The Second Objection

Now turning to the second question, why does Rawls exclude benevolent absolutism, burdened societies and outlaw states from the Society of Peoples? If these societies also have the

need for moral recognition and respect, denying them equal membership in the Society of Peoples would lead to bitterness and resentment. This would bring instability to the Society of Peoples. I argue that there is also a Rousseauvian response to this question. First, burdened societies and outlaw states do not respect the human rights of its members, and so denying them membership in the society of peoples is related to one of Rawls's aims of the law of peoples, that is, to meet the demand of amour-propre at least minimally for all persons. Second, liberal and decent peoples do not grant equal respect to benevolent absolutisms, burdened societies and outlaw states partly because they are not moral persons that have a secure sense of amour-propre.

One of the important reasons why Rawls says that outlaw states do not deserve respect is that these societies do not respect the human rights of their members, as Rawls writes: "An outlaw state that violates these rights is to be condemned and in grave cases may be subjected to forceful sanctions and even to intervention" (LP 81). But Charles Beitz (2000) argues that outlaw states and other societies should be included in the international original position:

> The exclusion of outlaw (and other) societies from the international original position means that no argument for human rights, made from the point of view of the original position, could establish that human rights have the kind of universality that is usually presumed. We want to say that people are entitled to be treated in certain ways (partly) in virtue of their characteristics as human beings and independently of considerations that might be particular to their own institutions and political cultures. (p.686)

Beitz argues that one of the reasons why outlaw states and other societies should be included in the international original position is that the people in these societies deserve equal respect regardless of their national identity. He argues that in the international original position,

if all persons in the outlaw states and other societies are included, parties who are free and equal rational persons would choose the global distributive principles of justice. Thus, Beitz argues that if Rawls wants to sincerely endorse the universality of human rights, he should start with the assumption that all persons deserve equal respect regardless of the origin of their nations.

In response to Beitz's objection, I argue that on the one hand the most urgent human rights are respected by all peoples in the society of peoples; on the other hand, burdened societies, outlaw states and benevolent absolutism are not moral persons and that they do not have a positive sense of amour-propre.

First, denying outlaw states and other societies membership in the law of peoples meets the demand of amour-propre. If outlaw states and other societies are included in the international original position, the most urgent human rights will not be respected by all peoples as they are not respected by these societies.[103] Denying them membership in the law of peoples, one the one hand, serves as an incentive for these societies to respect human rights; on the other hand, it also shows liberal and decent peoples' respect of human rights. Thus, Beitz is mistaken to say that the universality of human rights requires the inclusion of outlaw states and other societies in international original position; on the contrary, the universality of human rights requires the exclusion of outlaw states and other societies in the international original position.

Second, outlaw states and other societies are not moral persons and they do not have the positive sense of amour-propre, and so they should not be represented by the parties in the international original position. In other words, in the ideal theory of international justice, these

---

[103] This is, of course, true of outlaw states and burdened societies because they do not honor human rights. Rawls says that benevolent absolutism honors human rights, but why aren't they included in the society of peoples? David Reidy (2006, p.176) points out that the members of benevolent absolutism enjoy the content of human rights only through the good will of their ruler, not through their reciprocal commitment, and so they cannot "enjoy the content of their basic human rights as a matter of right."

societies are excluded because they are neither just nor decent. The burdened societies are not well-ordered because they suffer from unfavorable conditions. Rawls argues that burdened societies lack the necessary resources to become well-ordered and so liberal and decent peoples have the duty to assist burdened societies to become well-ordered. Once they become well-ordered, they will have a positive form of amour-propre. Thus, burdened societies will become members of the Society of peoples in due course. But outlaw states must not be tolerated by liberal and decent peoples. Rawls says: "Liberal and decent peoples have extremely good reasons for their attitude. Outlaw states are aggressive and dangerous; all peoples are safer and more secure if such states change, or are forced to change, their ways" (LP 81). Thus, parallel to Rawls's treatment of unreasonable persons, he argues that unreasonable societies are not to be respected by liberal and decent peoples. Indeed, outlaw states are like the slaveholder whose self-respect is enhanced by subjecting others to a lesser liberty. Outlaw states also have amour-propre. But their amour-propre is inflamed. They have a desire for domination, rather than a desire for equal standing with others in international relations. Thus, this desire of theirs must be suppressed by all reasonable (or not unreasonable) peoples. Unlike burdened societies and outlaw states, benevolent absolutism is not aggressive, and it respects the human rights of its members (not as a matter of right). But unlike liberal and decent peoples, the benevolent absolutism lacks the institutional structure needed to situate its members in various offices, etc., such that taken altogether they constitute one agent. As Rawls says, benevolent absolutism "does not give its members a meaningful role in making political decisions" (LP 92). Since benevolent absolutism doesn't incorporate its members into a single structure — it subjects them to the authority of its ruler, it is not one collective agent, but rather a ruler coordinating a

population. Thus, like burdened societies and outlaw states, benevolent absolutism ought not to be represented by the parties in the international original position.

4. Conclusion

I have argued that three themes of Rawls's law of peoples, that is, the two international original position, the idea of a realistic utopia, and the distinction between peoples and states, parallel, and are grounded in, Rousseau's concept of amour-propre. I have also argued that Rawls's argument that the law of peoples meets the demand of amour-propre, his stability argument and his arguments against political realism and distributive cosmopolitanism parallel, and are grounded in, Rousseau's concept of amour-propre. I have also considered two objections against Rawls's argument from the point of view of distributive cosmopolitanism. The first objection is that Rawls's law of peoples neglects women's need for recognition in decent hierarchical societies. The second objection is that Rawls's law of peoples neglects other societies' need for recognition such as benevolent absolutism, burdened societies and outlaw states. I have argued that there are Rousseauvian responses to both objections. To the first objection, I have argued that Rawls's treatment of decent peoples both satisfies the demand of amour-propre for decent peoples and the demand of amour-propre at least minimally for all persons. In response to the the second objection, I have argued that Rawls's treatment of outlaw states and other societies meets the demand of amour-propre for all persons. I have also argued that these societies are denied membership in the society of peoples because they are not moral persons that have a positive form of amour-propre. Therefore, there is a Rousseauvian interpretation of Rawls's law of peoples. In particular, Rousseau's concept of amour-propre plays a significant role in Rawls's law of peoples.

By recognizing the significance of amour-propre in Rawls's law of peoples, we can see that Rawls's law of peoples is not inconsistent with his theory of justice in the domestic case. This chapter thus also serves as a response to Rawls's critics who argues that Rawls's law of peoples is inconsistent with his justice as fairness. Furthermore, we have seen that Rousseau's concept of amour-propre plays a significant role in Rawls's stability argument in the law of peoples, and so this chapter also serves as the completion of my interpretation of Rawls's stability argument.

CHAPTER 6 CONCLUDING REMARKS

"Burdened as a man may be with the weight of tradition, he can yet prop open the gate of darkness with his shoulder to let the children through to the bright, wide-open spaces, to lead happy lives hence forward as rational beings."

(Lu Xun, "What Is Required of Us as Fathers Today")[104]

In the preceding chapters, I have defended the following theses. First, Rawls correctly interprets Rousseau's concept of amour-propre as having both a negative form and a positive form. The negative form of amour-propre is the desire for domination, whereas the positive form of amour-propre is the desire for having equal status with others. Second, for both Rawls and Rousseau, the desire for having equal status with others, that is, self-respect, rests on a person's equal citizenship and its fair value, not on his or her socio-economic status. Third, since Rawls's justice as fairness grants everyone equal citizenship and its fair value in the well-ordered society, justice as fairness satisfies the demand of amour-propre. I have also argued that since Rawls's justice as fairness meets the demand of amour-propre, people are motivated to act on and from the sense of justice in the well-ordered society of justice as fairness. I have argued that this explains why Rawls's justice as fairness is stable. Finally, I have argued that Rawls's law of

---

[104] The English translation is from *Lu Xun Selected Works, Vol. II*, Translated by Yang Xianyi and Galadys Yang, Beijing: Foreign Languages Press, 1980, p.57. Another English translation says: "Let the awakened man burden himself with the weight of tradition and shoulder up the gate of darkness. Let him give unimpeded passage to the children so that they may rush to the bright, wide-open spaces and lead happy lives henceforward as rational human beings" (Tsi-An Hsia, 1968, pp. 146-7). The original text is "自己背着因袭的重担，肩住了黑暗的闸门，放他们到宽阔光明的地方去；此后幸福的度日，合理的做人。"（鲁迅，"我们怎样做父亲"）

peoples meets the demand of amour-propre for both peoples and persons, and that this explains why Rawls's law of peoples is stable.

In the above argument, I have assumed that Rawls's argument that the legitimate bases of self-respect are equal basic rights and liberties and their substantive fair value, not equal socio-economic status is plausible. Given this assumption, I have argued that since Rawls's justice as fairness grants equal basic rights and liberties and their substantive fair value (secured by the second principle of justice) to everyone in the well-ordered society, justice as fairness meets the demand of amour-propre. Similarly, given this assumption, I have argued that since Rawls's law of peoples grants equal basic rights and liberties to all peoples and persons (in a minimal sense for the latter), the law of peoples meets the demand of amour-propre. In this chapter, I shall consider two objections against my assumption that the legitimate social bases of self-respect are equal basic rights and liberties and their substantive fair value. The first objection comes from some Chinese nationalists who argue that Chinese people's self-respect comes from their achievements and accomplishments in their history, especially the traditional culture of China.[105] The second objection is from luck egalitarians who argue that a person's self-respect comes from economic equality. I argue that these two objections are groundless.

1. Chinese Nationalists and Self-respect

Some Chinese Nationalists argue that Chinese people's self-respect rests on their achievements and accomplishments in their history, especially the traditional culture of China. I

[105] While I agree with Rawls that cultural achievements are often a proper bases of national pride, I think this is true only if the achievements constitute instances of moral development, for example, moving to decency and from decency to liberal democracy. This is the reason why I think that the traditional culture of China, which represents a morality of authority, cannot serve as the proper bases of national pride.

often hear people say that they are proud of being a Chinese because the tradition of old China has lasted for five thousand years, or because China was a great nation, or because their ancestors built the great wall, and so on. In recent years, I have also heard people say that they are proud of being a Chinese because China has the world's second largest economy. No doubt, some of these features are peculiar to China, but none of them, as I see it, contributes to a proper sense of self-respect. I will first study what these Chinese nationalists believe and why they hold this view. Then I will argue that the tradition of old China cannot be the legitimate base of self-respect.

1.1 Chinese Nationalists' View on Self-respect

Robert Weatherley (2014, p.20) claims that Chinese nationalists identify themselves with "the common myths, shared history and common public culture of their nation." The common myths refer to the greatness of China as a nation. The shared history refers to the achievements and accomplishments in their history. The common public culture refers to the Confucianism tradition of old China, the common language and writing system and the four great Chinese inventions: porcelain, paper, printing and gunpowder. He argues that Chinese nationalism was a powerful weapon against western colonialism in the early twenty-century, but it was also an obstacle for China's democratization. Robert Weatherley (2014) points out that Sun Yat-sen's Republic of China was a constitutional republic that had a constitution and a parliament. But he argues that democracy did not lead China toward political independence from western imperialism.[106] So Sun and Chiang Kai-shek believed that the mass of Chinese people were not

[106] He Baogang (2004, p.176) argues that democracy did not lead to national unity in China in the early twenty-century due to the rise of regional militarists and that this forced both leaders in the Kuomintang Party and the Chinese Communist Party to shift from democracy to authoritarianism.

ready for a parliamentary democracy. China shifted towards authoritarianism. Individual rights and liberties were sacrificed in order to expel foreign enemies from China. After the establishment of People's Republic of China (PRC), even though the Japanese Army was expelled out of the land of China, protecting China's independence from external threats remains a primary task of the PRC. Some of the basic human rights are not honored by the PRC because human rights talk is taken to be a form of western imperialism.[107]

Yingjie Guo (2004, p.17) distinguishes state nationalism and cultural nationalism in today's China. State nationalism, advocated by Chinese government officials, "portrays the state [or the communist party] as the embodiment of the nation's will, seeking for its goal the kind of loyalty and support granted the nation itself and trying to create a sense of nationhood among all its citizens" (p.17). Cultural nationalism, advocated by Chinese intellectuals, tries to recreate "a widespread awareness of the myths, history, and linguistic traditions of the community" (p.18). While state nationalism and cultural nationalism often go hand in hand, cultural nationalists do not identify themselves with advocates of the party-state. They also think that the socialists' and Marxists' elements in the party-state are illegitimate. But both state nationalism and cultural nationalism think that human rights, liberty and democracy are parts of western culture that contribute to Western imperialism. State nationalism argues that it brings harm to the party-state to demand human rights, liberty and democracy (cf. p.29). Cultural nationalism thinks that human rights, liberty and democracy are parts of colonial culture that are detrimental to national self-esteem (cf. p.112).

[107] Also, an important ideological reason is that socialism is regarded as the only correct path toward equality and freedom.

We can see that Chinese nationalists think that equal basic liberties and rights are not the legitimate bases of self-respect for Chinese people for two reasons. First, they think that human rights, liberty and democracy are necessarily linked with western imperialism and colonialism. Second, they think that they are foreign to us because they are not a part of the tradition of old China. They think the legitimate bases of self-respect for Chinese people must, and should, originate from their own culture. And so since human rights, liberty and democracy are not in the tradition of old China, they cannot be the legitimate bases of self-respect for Chinese people. Thus, they think that China should not advance human rights, liberty and democracy to secure Chinese people's self-respect; rather, they think that China must make the tradition of old China universal, and make other peoples recognize the value of the traditional Chinese culture.

1.2 Why Can the Tradition of Old China not be the Proper Bases of Self-respect?

I think that it is understandable that Chinese nationalists psychologically connect western imperialism with human rights, liberty and democracy. This psychological connection stems partly from the pain and miseries that Chinese people experienced in their struggle for national independence, and partly from their failure in achieving national independence by democratization. But even though in the hearts of Chinese nationalists, there is a psychological connection that links western imperialism with human rights, liberty and democracy, there is no necessary conceptual connection between them. In fact, if Rawls is right, and I think he is, liberal peoples have a positive sense of amour-propre that inclines them to respect other nations' rights, which include their right of non-intervention unless that nation severely violates basic human rights. Thus, liberal democracies necessarily do not have a tendency to invade other countries.

The western countries that invaded China in the last century were not liberal democracies by then.

Chinese nationalists may argue that it does not matter whether liberal democracies are aggressive or not. They may argue that what matters is that liberal democracy is a foreign thing for Chinese people. And so equal basic rights and liberties and their substantive value cannot be the rightful bases of self-respect for Chinese people. They argue that the legitimate bases of self-respect for Chinese people are the tradition of old China. In what follows, I will give two case studies and argue that the tradition of old China cannot be the proper bases of self-respect because "filial piety" and "chastity," which are two core values of the traditional culture of China, are corrosive of self-respect. I will also argue that human rights, liberty and democracy are a part of, and therefore are not foreign to, the domestic public culture of China because they are embodied in the work of Lu Xun and other Chinese liberals.

1.2.1 Two Case Studies: "Filial Piety" and "Chastity" are Corrosive of Self-respect

"Filial piety" and "chastity" are two of the core values of the traditional culture of China because they regulate family relations which mirror the relations between the rulers and the subjects. "Filial piety" requires children to treat their parents with "piety" because their parents give birth to them and raise them. According to the tradition of old China, children do not have self-ownership; rather they are owned by (or belong to) their parents. "Chastity" requires women to treat their husbands with "chastity," that is, they do not have the right to remarry after the deaths of their husbands and they cannot have relations with other men. In other words, women do not have self-ownership too; rather they are owned by (or belong to) their husbands. These two virtues were highly praised by intellectuals and moralists in the old China. They are two of

the three Cardinal Guides (三纲) that regulate the relations between people. They are also regarded as the primary motivating force for the stability of the old China. But as we will see, these two virtues are corrosive of a person's self-respect. And so they are not the proper motivating force for stability.

In the Beginning of *The Classic of Filial Piety (The Hsiao King)*, the Master (Confucius) asked his student:

> The ancient kings had a perfect virtue and all-embracing rule of conduct, through which they were in accord with all under heaven. By the practice of it the people were brought to live in peace and harmony, and there was no ill-will between superiors and inferiors. Do you know what it was? (*The Hsiao King*, I)[108]

His student, Zeng, did not know the answer. So the Master answered:

> (It was filial piety.) Now filial piety is the root of (all) virtue, and (the stem) out of which grows (all moral) teaching. Sit down again, and I will explain the subject to you. Our bodies- to every hair and bit of skin- are received by us from our parents, and we must not presume to injure or wound them. This is the beginning of filial piety. When we have established our character by the practice of the (filial) course, so as to make our name famous in future ages and thereby glorify our parents, this is the end of filial piety. It commences with the service of parents; it proceeds to the service of the ruler; it is completed by the establishment of character. (ibid.)

---

[108] the English translation is from *The Sacred Books of the East: The Texts of Confucianism,* vol. III, part I: The Shu King, The Religious Portions of the Shih King, The Hsiao King, translated by James Legge, Oxford, Clarendon Press, 1879, pp.465-6.

Here it says that filial piety is the origin of all virtues and the essence of moral teaching. It says that with the practice of filial piety, the stability (peace and harmony) of the society is achieved. The asymmetrical emphasis on filial piety from children to parents is characteristic of the tradition of the old China. Filial piety advocates the self-sacrificing spirit of children to please their parents. *The Twenty-four Paragons of Filial Piety* manifests this spirit in its extreme. For example, among others, there is the story of Dong Yong who enters servitude to pay for his father's Funeral and the story of Wang Xiang who lies down on the ice to get carp for his stepmother.[109] In these stories, filial piety is praised as the virtue of highest priority that outweighs other virtues and human goods. And so Dong Yong is praised as a "filial son" because he sells himself as a slave to bury his father's body despite the fact that servitude is corrosive of his own self-respect. The story of Wang Xiang also shows how filial piety is corrosive of self-respect. Wang Xiang's stepmother had always treated him badly. But Wang Xiang was a "filial son." So he practiced his duty as a son, lying on the ice to get carp for her stepmother because she had a craving for carp in her sickness. This story illustrates how the subjects in the old China are taught to treat their rulers. Their rulers, like Wang Xiang's stepmother, can treat their subjects badly, but the subjects have the duty to do as the rulers command. "Filial piety" does not affirm a person's self-respect; rather, it nurtures the evils of servitude that "stabilize" the hierarchical society.

"Chastity," similar to "filial piety," emphasizes the asymmetrical relation between the rulers and their subjects. "Chastity" requires women to treat their husbands as their masters, and

---

[109] These stories are not the cruelest stories in *The Twenty-four Paragons of Filial Piety*. They are selected simply because they illustrate how filial piety is corrosive of self-respect. The cruelest story among them is the story of Guo Ju who tries to bury his son in order to save food for his mother. Lu Xun feels this story repugnant. This story is not selected in today's popular versions of *The Twenty-four Paragons of Filial Piety*.

so men can have many concubines but women must keep their "chastity" even after the death of their husband. "Chastity" also undermines a person's self-respect. It enhances women's servitude that "stabilizes" the patriarchal society.

1.2.2 Lu Xun's Criticisms of "Filial piety" and "Chastity"

After the May-fourth movement in 1919, Lu Xun wrote that the reform of China should start with the reform of family, and that the reform of family should start with the abandonment of the old tradition of "filial piety" and "chastity." He argues that the tradition of old China is the reverse of the natural order:

> The young should take first place, but instead it is taken by the old. The emphasis should be on the future, but instead it is on the past. The elder generation is sacrificed to the generation before it; yet, with no means to outlive itself, it expects the younger generation to sacrifice itself for the elder's sake, destroying all that could carry it forward. (Lu Xun, 1980, p.60)

Lu Xun argues that the tradition of old China assumes that "the young should be sacrificed entirely to the old" (ibid.). He argues that this assumption is at variance with the natural order. He argues that the traditions of old China of "filial piety" and "chastity" should give way to European and American family tradition, that is, "the young and the weak usually take first place" (ibid., p.61). He argues that parents should follow the natural way, that is, to love their children first, rather than ask for their sacrifice first. He argues that this is the path toward the emancipation of the young, and toward their happy lives.

Lu Xun's thoughts are deeply influenced by Rousseau. In an essay entitled "Rousseau and Personal Taste," he tries to defend Rousseau's ideas of freedom and equality, especially, the

equal rights for men and women,[110] against Rousseau's critics (cf. p.377). His objection against "chastity" is in the same vein.[111] His thought that love should be the root of morality also parallels Rousseau's thought that moral sentiments are derived from natural sentiments. But unfortunately, while Lu Xun is the most influential figure in modern Chinese literature and a source of the domestic public culture of China, his thoughts on moral education and on political liberties are underestimated.[112]

Since the emphasis on freedom and equality is not absent from the work of Lu Xun's and other Chinese liberals,[113] which have been regarded as an important source of the domestic public culture of China, they are not foreign to us. Furthermore, Chinese nationalists cannot say that China does not have the soil to bring liberal democracy into fruition because, as Lu Xun, Rousseau and Rawls all point out, the first step toward a society of mutual respect and self-esteem is natural love of which all human beings are capable.

To sum up, I have argued that some Chinese nationalists are mistaken to connect western imperialism with human rights, liberty and democracy. I have also argued that Chinese nationalists cannot ground their self-respect on the tradition of old China because "filial piety" and "chastity," two of the core values of the traditional culture of China, are corrosive of self-

---

[110] The inequality of women in the villages of China is exhibited in the organization of the village council, the members of which are elected by villagers. In the village council, there is only one single woman called "women's representative," whose primary task, over the past 30 years, has been birth control (cf. Manoranjan Mohanty and Mark Selden, 2007, p.465).

[111] Lu Xun refutes the spiritualists' claim that women's chastity is the remedy for the moral degeneration in China in the early twenty-century. He has the deepest compassion for women's miseries in the old China where they have no right to remarry after the death of their husband. See *Lu Xun Selected Works*, Vol. II, 1980, pp.13-25.

[112] Lu Xun was praised by Mao Tse-tung as "the most correct, the bravest, the firmest, the most loyal and most zealous hero who stormed and broke up the enemy's front" (Tsi-An Hsia, 1968, p.102). He was one of the leaders of the Chinese League of Leftist Writers which aimed to support the activities of the Communist Party and was supported by it from 1930-1935. But according to Tsi-An Hsia, from 1935 to the time of his death, Lu Xun "had lost much of his interest in the League and had been sorely disappointed in it at times" (ibid., p.106).

[113] For example, Hu Shih argues that Chinese people's self-confidence cannot rest on the past glory of our ancestors; rather he argues that it should rest on the reflection about what has gone wrong in the past (cf. Hu Shih, 1934).

respect. I have also argued that the value of liberty and equality has been recognized by Chinese intellectuals such as Lu Xun whose works are a part of the domestic public culture of China. And so Chinese nationalists cannot say that human rights, liberty and democracy are foreign to us. More importantly, they are not foreign to us because as Lu Xun, Rousseau and Rawls all point out, we are all capable of natural love which is the first step toward a society of mutual respect and self-esteem (or a liberal democracy).

2. Luck Egalitarianism's Objections against Rawls[114]

In Chapter 3, I argued that the public affirmation of the two principle of justice, especially the public affirmation of equal citizenship, ensures the social bases of self-respect for everyone in the well-ordered society of justice as fairness. I have argued that the legitimate social bases of self-respect are equal citizenship, not equal socio-economic status. I have argued that Rawls's difference principle, that is, social and economic inequalities are to be the greatest benefit of the least advantaged people, is partly used for securing equal basic rights and liberties for everyone substantively. And so the difference principle also lends support for a person's self-respect. But one might wonder why the difference principle is required to affirm a person's self-respect, not other principles such as G.A Cohen's strict reading of the difference principle. In what follows, I will first introduce Joshua Cohen's argument from self-respect for Rawls's difference principle. I

[114] Here I mainly consider the form of luck egalitarianism that G. A Cohen proposes in *Rescuing Justice and Equality*. Luck egalitarianism maintains that inequalities are unjust if they are the results of brute luck. Cohen says that "The view that accidental inequality is unjust is suggested in certain passages in Ronald Dworkin's 'Equality of Resources.' Dworkin did not himself go so far as to endorse the view in its full generality, but Richard Arneson and I proposed its generalization, and Elizabeth Anderson aptly dubbed the proposal 'luck egalitarianism'" (Cohen 2008, p.8). Rawls thinks that luck egalitarianism mistakenly ties a person's self-respect with an ideal of cosmic order — being immunized from brute luck.

will then introduce G.A Cohen's objection against Joshua Cohen's argument. In the end, I will argue against G. A Cohen's objection.

2.1 Joshua Cohen's Argument from Self-respect for Rawls's Difference Principle

Joshua Cohen argues that Rawls's difference principle lends greater support for self-respect than alternative principles. His argument goes like this: (1) "the value of liberties to a person depends on the level of resources available to the person" (Joshua Cohen, 1989, p.738); and (2) Rawls's difference principle provides greater minimum value of the liberties than alternative principles; and (3) self-respect depends on the value of the liberties "since a greater value enables a more confident pursuit of one's aims" (ibid., p.739); and so justice as fairness lends greater support for self-respect than alternative principles.

He also argues that the public affirmation of the fair value of liberty, or as he puts it, "the maximal worth of the liberties" (ibid.), provides the recognitional bases for self-respect. Once citizens know, and know that others know, that their worth does not rest on social advantages such as social positions, natural endowment and good fortune, they will fully affirm others' worth and express respect for others. Thus, he argues that Rawls's difference principle provides greater support for self-respect than alternative principles.

2.2 G. A Cohen's Objection Against Rawls's Difference Principle

G. A Cohen argues that Rawls's difference principle, or what he calls "the lax reading of the difference principle," faces the same objection that Joshua Cohen raises (cf. G. A Cohen 2008, p.77). Thus, according to G. A Cohen, since his strict reading of the difference principle provides greater minimum value of the liberties than Rawls's difference principle, his strict

reading of the difference principle lends greater support to self-respect than Rawls's difference principle.[115] Furthermore, he argues that Rawls's difference principle does not really commit to treating every person as equals regardless of their social position, natural endowment and good fortune. He argues that Rawls's difference principle is not a basic principle of justice but a compromise to the fact that the talented and the rich need unequal incentives to work hard. He argues that this fact is only a fact when the talented and the rich make it true. He argues that if the talented and the rich, as Rawls describes in the well-ordered society of justice as fairness, have the sense of justice, they would not need unequal incentives to work hard to advance the interests of the rest. And so he argues that his strict reading of the difference principle which "counts inequalities as necessary only when they are, strictly, necessary, necessary, that is, apart from people's chosen intentions" is a basic principle of justice (ibid., p.69).

G. A Cohen illustrates his point with the example of two brothers, A, and B, who try to decide to move from New York to Chicago or Boston. Suppose A and B are at the benefit levels of 6 and 5 in New York. If they move to Chicago, their levels will rise to 10 and 5.1. If they move to Boston, their levels will rise to 8 and 7. G. A Cohen argues that according to Rawls's difference principle, it is reasonable for A and B to move to Chicago even though the economic inequalities are enlarged. He argues that for Rawls, these inequalities are justified because the prospect of the less advantaged (B) in Chicago is better than in New York. If we take the factual premise that A does not have unequal incentives to move to Boston into account, Rawls's difference principle, as Joshua Cohen suggests, provides the maximal minimum value of the

[115] Richard Penny (2013) develops G. A Cohen's argument from self-respect against Rawls's difference principle. He argues that Rawls's difference principle undermines a person's self-confidence because it leaves "those who are least advantaged with a smaller absolute bundle of resources with which to form and pursue their plan of life" (p.346). He also argues that Rawls's difference principle undermines the value of mutual beneficence, mutual respect and fraternity. For an objection against Penny's view, see Harrison P. Frye (2015).

liberties. But since the factual premise would not be true if A values fraternity and decides to move to Boston, G. A Cohen argues that Rawls's difference principle does not provide greater minimal value of the liberties than his strict reading of the difference principle. And so he argues that his strict reading of the difference principle lends greater support to self-respect than Rawls's difference principle.

The proponents of Rawls might argue that Rawls's two principles of justice only apply to the basic structure of the well-ordered society of justice as fairness, not to personal choices. They might argue that insofar as citizens in the well-ordered societies comply with the rules of the just basic structure, personal choices do not affect distributive justice (cf. ibid., p.126). In particular, they might argue that even though the rich and the talented are "market maximizers," citizens in the well-ordered society of justice as fairness are still granted the social bases of self-respect secured by the just basic structure. But G. A Cohen argues that this basic structure objection faces two difficulties. The first difficulty is that restricting justice to the basic structure undermines the realization of fraternity, self-respect and the sense of justice. The second difficulty is that Rawls's idea of basic structure faces a dilemma: either the basic structure only includes the coercive structure, then this restriction is arbitrary; or it also includes social practices and personal choices, then this restriction collapses (cf. ibid., p.137). In particular, he argues that the reason why restricting justice to the basic structure undermines self-respect of the worse off is that their positions are made inferior to others by the basic structure by acquiescing the rich and the talented people's needs for unequal incentives. Thus, he concludes that the basic structure objection fails.

2.3 My Argument Against G. A Cohen's Objection

A number of commentators have provided Rousseauvian responses to G. A Cohen's objection.[116] For example, Robert Jubb (2011, p.257) argues that G. A Cohen's strict reading of the difference principle "fails to generate its own support because the policies needed to realize it are ones which tend to demean and disfigure the lives of some of those who live under them." Jeffrey Bercuson (2014) argues that G. A Cohen's strict reading fails to take men as they are and political institutions as they might be (cf. pp.140-2). My argument against G. A Cohen's objection is in the same vein. I argue that G. A Cohen's strict reading of the difference principle is based on an unrealistic moral psychology.

G. A Cohen argues that Rawls's incentive argument for the difference principle is based on the assumption that human beings are unjustly selfish. He says:

> No explanation was given in 1971 of why the incentives are necessary to the stated ends, but it follows from my argument against Rawls's position on incentives that, if unequalizing incentives are truly necessary from the point of view of the interest of the badly off, then they are necessary only because of an infirmity in human nature, to wit, a certain unjust selfishness, that is more or less acknowledge in the "Justice as Fairness" passage but that gains no mention, as such, in the corresponding *Theory* passage. (G. A Cohen, 2008, p.178)

As I have said, for G. A Cohen, Rawls's difference principle is a compromise to the fact that the rich and the talented need unequal incentives to work hard. This fact, he argues, is made

---

[116] There are many other objections against G. A Cohen's view. For example, David Miller (2008), Thomas Pogge (2008), Robert Jubb (2009), Jan Narveson (2010) and Lea Ypi (2012) argue that Cohen's claim that justice is fact-insensitive is implausible or insignificant. Andrew Williams (1998) argues that Cohen's strict reading of the difference principle fails to satisfy the principle of publicity. Alan Thomas (2011) argues that Cohen's critique of Rawls's special incentives is unsustainable.

true intentionally by the rich and the talented. And so if the incentive argument is voiced from the rich and the talented, the poor and the untalented cannot reasonably accept it. Here G. A Cohen argues that Rawls's incentive argument is based on the assumption that human beings are unjustly selfish. For if this assumption is false, Rawls's factual premise that the rich and the talented need unequal incentives to work hard will not be true. G. A Cohen thus argues that Rawls's difference principle is not a basic principle of justice and that it undermines the self-respect of the worse-off people. His response to the basic structure objection also relies on this objection against Rawls's moral psychology.

One of the cruxes of the debate between Rawls (and Joshua Cohen) and G. A Cohen is whether Rawls's incentive argument is based on the assumption that human beings are unjustly selfish. Rawls, of course, has never said that human beings are unjustly selfish. He assumes that moral persons are free and equal, rational and reasonable. He argues that human beings have the capacity to form, revise and pursue their own conception of the good, and that they have the capacity to act on and from the sense of justice. G. A Cohen argues that Rawls has an incoherent account of human nature because Rawls argues that human beings have both the capacity for the sense of justice and the capacity for the conception of the good. For Cohen, the citizens who act from Rawls's difference principle are not really just; rather, he thinks that they are unjustly selfish.

But as we have seen in Chapter 4 Rawls has a realistic moral psychology. For Rawls and Rousseau, the moral sentiments of the sense of justice are derived from the natural sentiments of self-love. In the first stage of moral development, if children see that their parents manifestly love them, they will love their parents back. This returning sentiment of natural love is the first step toward the formation of the sense of justice. In the second stage of moral development,

given that a person has acquired natural love in the first stage, if he or she sees that his or her associates fulfill the duties and obligations set up by the association they belong to with evident intention, he or she will have feelings of friendship toward them. In the third stage of moral development, given that a person has experienced the first two stages, if he or she sees that the just institution is beneficial to him and those around him, he or she will acquire the sense of justice. In each stage of moral development, the reciprocity of disposition is exhibited. But in the second stage, the reciprocity of disposition is more elevated than in the first stage. And in the third stage, the reciprocity of disposition is the most elevated. Once a person has acquired the sense of justice, he or she will act on and from the sense of justice regardless of natural or social contingencies. And so, G.A Cohen is mistaken to say that Rawls's moral persons are unjustly selfish. Rather, they are simply "just." For Rawls and Rousseau, the sense of justice is not incompatible with self-love or self-interest. On the contrary, the sense of justice originates from self-love or self-interest. As Rousseau says:

> Our first duties are to ourselves; our primary sentiments are centered on ourselves; all our natural movements relate in the first instance to our preservation and our well-being. Thus, the first sentiment of justice does not come to us from the justice we owe but from that which is owed us; and it is again one of the mistakes of ordinary educations that, speaking at first to children of their duties, never of their rights. (E 97)

For Rousseau, Cohen's insistence on the strict reading of the difference principle is but "again one of the mistakes of ordinary educations that, speaking at first to children of their duties, never of their rights." Cohen's strict reading of the difference principle requires people to sacrifice their basic interests to advance the interests of others. Of course, it is not always

unreasonable to expect others to sacrifice their less important interests to advance my important interests. But it is less clear whether I have a duty to give you ten dollars to make us equal if I have w dollars and if you have w-20 dollars. If, as Cohen suggests, my giving you ten dollars enhances your self-respect, then the self-respect that you have is a fragile self-respect that can easily be broken if I refuse to give you the ten dollars. Self-respect should have a stronger foundation, one that is not easily broken by people's intentions. In any case, if the society well-ordered by the strict reading of the difference principle does not give me the freedom to keep my ten dollars, I will hesitate to live in that society. I hesitate to live in that society not because I value ten dollars, but because I value my freedom to keep my ten dollars. If Cohen's society violates this freedom of mine, it violates "the first sentiment of justice." Thus, Cohen's society may not be a just society in which equal basic liberties and rights are secured to affirm the self-respect of everyone. Rather, some of these liberties may be violated in order to enhance economic equality on the strict reading of the difference principle. Thus, Cohen's strict reading of the difference principle cannot provide the social bases of self-respect. And so it fails to meet the demand of amour-propre.

Thus Joshua Cohen is correct that Rawls's difference principle provides greater minimum value of the liberties than alternative principles. It lends greater support to self-respect than Cohen's strict reading of the difference principle because Cohen's principle fails to meet the demand of amour-propre. The reason why Cohen's principle fails to meet the demand of amour-propre is that it assumes an unrealistic moral psychology, according to which, people ought to sacrifice their basic interests to advance the interests of others.

3. Conclusion

In this chapter, I have defended Rawls's argument that the legitimate social bases of self-respect are equal basic liberties and rights and their substantive fair value, not socio-economic status. I have argued that some Chinese nationalists mistakenly ground Chinese people's self-respect on the tradition of old China. I have argued that the tradition of old China cannot be the proper bases of self-respect because "filial piety" and "chastity," the two core values of the traditional culture of China, are corrosive of self-respect. I have also argued that equal basic liberties and rights and their substantive fair value can be the proper bases of self-respect for Chinese people because they are not foreign to us — they are evident in the writings of Lu Xun and other Chinese liberals which fuse into the domestic public culture of today's China.

I have also argued that luck egalitarians mistakenly ground self-respect in economic equality. I have argued that economic equality based on Cohen's strict reading of the difference principle cannot be the proper bases of self-respect because Cohen's principle fails to provide the social bases of self-respect. I have argued that the reason why it fails to provide the social bases of self-respect is that it assumes an unrealistic moral psychology, according to which, people must sacrifice their freedom to advance the greater economic interests of others.

BIBLIOGRAPHY

Arneson, Richard J. 2000. “Luck Egalitarianism and Prioritarianism”. *Ethics* 110 (2):339-349.

Barry, Brian M. 1977. “Rawls on Average and Total Utility: A Comment.” *Philosophical Studies* 31 (5):317 - 325.

Barry, Brian M. 1995. *Justice as Impartiality*. Oxford University Press.

Baogang He. 2004. “China’s National Identity: A Source of Conflict between Democracy and State Nationalism” in *Nationalism, Democracy and National Integration in China*, edited by Leong H. Liew and Shaoguang Wang, 170-195. London and New York: RoutledgeCurzon.

Baumrin, Bernard H. 1976. “Autonomy in Rawls and Kant.” *Midwest Studies in Philosophy* 1 (1):55-57.

Beitz, Charles R. 2000. “Rawls's Law of Peoples.” *Ethics* 110 (4):669-696.

Bercuson, Jeffrey. 2014. *John Rawls and the History of Political Thought: The Rousseauvian and Hegelian Heritage of Justice as Fairness*. London: Routledge, Taylor & Francis Group.

Bertram, Christopher. 2004. *Routledge Philosophy Guidebook to Rousseau and the Social Contract*. London: Routledge.

Bird, Colin. 2010. “Self-respect and the Respect of Others.” *European Journal of Philosophy* 18 (1):17-40.

Bix, Brian. 2004. *A Dictionary of Legal Theory*. Oxford University Press.

Bloom, Allan. 1975. “Justice: John Rawls Vs. The Tradition of Political Philosophy.” *The American Political Science Review*, 69 (2): 648–662.

Bloom, Allan. 1979. “Introduction,” in *Emile: Or, on Education*. New York: Basic Books.

Braybrooke, David. 1973. “Utilitarianism with a Difference: Rawls's Position in Ethics.” *Canadian Journal of Philosophy* 3 (2):303 - 331.

Brock, Gillian. 2010. “Recent Work on Rawls’s Law of Peoples: Critics vs Defenders”. *American Philosophical Quarterly* 47 (1):85-101.

Brooke, Christopher. 2012. *Philosophic Pride: Stoicism and Political Thought from Lipsius to Rousseau*. Princeton: Princeton University Press.

Buchanan, Allen E. 1990. “Justice as Reciprocity versus Subject-centered Justice.” *Philosophy and Public Affairs* 19 (3):227-252.

Buchanan, Allen E. 2000. “Rawls’s Law of Peoples: Rules for a Vanished Westphalian World.” *Ethics* 110 (4):697-721.

Buchanan, Allen E. 2004. *Justice, Legitimacy, and Self-Determination: Moral Foundations for International Law*. Oxford University Press.

Buchanan, Allen E. 2006. “Taking the Human out of Human Rights” in *Rawls's Law of Peoples: A Realistic Utopia?*, edited by Rex Martin & David Reidy, 150-168. Blackwell.

Button, Mark E. 2008. *Contract, Culture, and Citizenship: Transformative Liberalism from Hobbes to Rawls*. University Park, Pa: Pennsylvania State University Press.

Cabrera, Luis. 2001. “Toleration and Tyranny in Rawls’s ‘Law of Peoples.’” *Polity*, 34 (2): 163–179.

Carson, Thomas L. 1983. “Strict Compliance and Rawls's Critique of Utilitarianism.” *Theoria* 49 (3):142-158.

Cassirer, Ernst. 1954. *The Question of Jean-Jacques Rousseau*. New York: Columbia University Press.

Chazan, Pauline. 1993. “Rousseau As Psycho-Social Moralist: The Distinction between Amour De Soi and Amour-Propre.” *History of Philosophy Quarterly* 10.4: 341-354.

Cohen, G. A. 2008. *Rescuing Justice and Equality*. Cambridge, Mass: Harvard University Press.

Cohen, G. A. 1995. *Self-Ownership, Freedom, and Equality*. Cambridge University Press.

Cohen, Joshua. 1989. "Democratic equality." *Ethics* 99 (4):727-751.

Cohen, Joshua. 1997. "The natural goodness of humanity." in, *Reclaiming the History of Ethics: Essays for John Rawls*, edited by Andrews Reath, Barbara Herman, Christine M. Korsgaard & John Rawls, 102-39. Cambridge: Cambridge University Press.

Cobban, Alfred. 1934. *Rousseau and the Modern State.* London: G. Allen & Unwin Ltd.

Costa, M. Victoria. 2010. *Rawls, Citizenship, and Education*. Routledge.

Crocker, Lester G. 1968. *Rousseau's Social Contract: An Interpretive Essay. Cleveland: Press of Case Western Reserve University*.

Cureton, Adam. 2013. "From Self-Respect to Respect for Others." *Pacific Philosophical Quarterly* 94 (2):166-187.

Darwall, Stephen L. 1976. "A Defense of the Kantian Interpretation." *Ethics* 86 (2):164-170.

Darwall, Stephen L. 1980. "Is There a Kantian Foundation for Rawlsian Justice?" in *John Rawls' Theory of Social Justice*, edited by H.G. Blocker & E. Smith. Athens. OH: Ohio University Press.

Darwall, Stephen L. 1977. "Two Kinds of Respect." *Ethics* 88 (1):36-49.

De Francisco, Andrés. 2006. "A Republican Interpretation of the Late Rawls." *Journal of Political Philosophy* 14 (3):270–288.

DeArmitt, Pleshette. 2013. *The Right to Narcissism: A Case for an Im-Possible Self-Love*.

Deigh, John. 1996. *The Sources of Moral Agency: Essays in Moral Psychology and Freudian Theory*. Cambridge University Press.

Dent, N J. H. 1988. *Rousseau: An Introduction to His Psychological, Social and Political Theory*. Oxford: Blackwell.

Dent, N J. H. 1992. *A Rousseau Dictionary*. Oxford: Blackwell Reference.

Deranty, Jean-Philippe. 2009. *Beyond Communication: A Critical Study of Axel Honneth's Social Philosophy*. Brill.

Estlund, David. 1996. "The Survival of Egalitarian Justice in John Rawls's *Political Liberalism*" *Journal of Political Philosophy*, 4(1): 68—78, reprinted in Kukathas, Chandran (ed.) *John Rawls: Critical assessments of leading political philosophers; Volume IV:* ***Political Liberalism*** *and* ***The Law of Peoples***, 380—391, Routledge, London.

Eyal, Nir. 2005. "'Perhaps the Most Important Primary Good': Self-respect and Rawls's Principles of Justice." *Politics, Philosophy and Economics* 4 (2):195-219.

Freeman, Samuel. 2003. "John Rawls–an Overview." In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 1-59. Cambridge University Press.

Freeman, Samuel. 2007. *Rawls*. Routledge.

Frye, Harrison P. 2015. "Putting Incentives in Context: A Reply to Penny." *Res Publica* 21 (1):93-98.

Garthoff, Jon. 2013. "Reconstructing Rawls: The Kantian Foundations of Justice as Fairness. By Robert S.Taylor. University Park: The Pennsylvania State University Press, 2011. 360p. $74.95." *Perspective on Politics*, 11(1):285-6.

Garthoff, Jon. 2016. "Rawlsian Stability." *Res Publica* 22 (3):285-299.

Gauthier, David P. 2006. Rousseau: *The Sentiment of Existence*. Cambridge: Cambridge University Press.

Gaus, Gerald. 2013. "The Turn to a Political Liberalism." In *A Companion to Rawls*, edited by Jon Mandle & David Reidy, 235-250. Hoboken: Wiley-Blackwell.

Guo, Yingjie. 2004. *Cultural Nationalism in Contemporary China: The Search for National Identity Under Reform*. London: RoutledgeCurzon.

Guyer, Paul. 2000. *Kant on Freedom, Law, and Happiness*. Cambridge: Cambridge University Press.

Hayfa, Tarek. 2004. "The Idea of Public justification in Rawls's Law of Peoples." *Res Publica* 10 (3):233-246.

Hedrick, Todd. 2015. "The History of Political Thought: The Rousseauvian and Hegelian Heritage of Justice as Fairness, By Jeffrey Bercuson, Routledge, 2014. Pp. xiii + 150. ISBN 978-0-415-73739-5. $140.00 (hbk)", *International Journal of Philosophical Studies*, 23(2): 296-301.

Held, David, and Maffettone, Pietro. 2016. "Introduction". in *Global Political Theory*, edited by David Held & Pietro Maffettone, 1-21. Malden, MA: Polity Press.

Hill, Thomas E. 1991. *Autonomy and Self-Respect*. Cambridge University Press.

Hill, Thomas E. 2014. "Stability, A Sense of Justice and Self-respect", in *A Companion to Rawls*, edited by Jon Mandle & David Reidy, 200-215. Hoboken: Wiley-Blackwell.

Honig, Bonnie. 1993. *Political Theory and the Displacement of Politics*. Ithaca: Cornell University Press.

Huemer, Michael. 1996. "Rawls's Problem of Stability." *Social Theory and Practice* 22 (3):375-395.

Lu, Xun. 1980. *Lu Xun: Selected Works Volume Two*. Beijing: Foreign Languages Press.

Hsia, Tsi-An. 1968. A. *The Gate of Darkness: Studies on the Leftistliterary Movement in China*. University of Washington Press.

James, Aaron. 2005. “Constructing Justice for Existing Practice: Rawls and the Status Quo.” *Philosophy and Public Affairs* 33 (3):281-316.

James, Aaron. 2016. “Fairness in Trade”. in *Global Political Theory*, edited by David Held & Pietro Maffettone, 213-235. Malden, MA: Polity Press.

James, David. 2011. “Rousseau on Needs, Language and Pity: The Limits of ‘Public Reason’.” *European Journal of Political Theory* 10 (3):372-393.

Jeanne S. Zaino. 1998. “Self-respect and Rawlsian Justice.” *The Journal of Politics*, 60 (3):737-753.

Johnson, Oliver A. 1974. “The Kantian Interpretation.” *Ethics* 85 (1):58-66.

Johnson, Oliver A. 1977. “Autonomy in Kant and Rawls: A reply.” *Ethics* 87 (3):251-254.

Jubb, Robert. 2009. “Logical and Epistemic Foundationalism about Grounding: the Triviality of Facts and Principles.” *Res Publica* 15: 337–353.

Jubb, Robert. 2011. “Rawls and Rousseau: Amour-Propre and the Strains of Commitment.” *Res Publica* 17 (3):245-260.

Kant, Immanuel, Theodore M. Greene, Hoyt H. Hudson, and John Silber. 1960. *Religion Within the Limits of Reason Alone*. New York: Harper & Row.

Kant, Immanuel, 1996. *Practical Philosophy*, translated and edited by Mary J. Gregor, in *The Cambridge Edition of the Works of Immanuel Kant*, edited by Paul Guyer and Allen W. Wood. Cambridge University Press.

Kaufman, Alexander. 2012. “Rawls and Kantian Constructivism.” *Kantian Review* 17 (2):227-256.

Kavka, Gregory S. 1975. “Rawls on Average and Total Utility.” *Philosophical Studies* 27 (4):237 - 253.

Keat, Russell and Miller, David. 1974. "Understanding Justice." *Political Theory* 2 (1):3-31.

Kim, H. 2015. "A Stability Interpretation of Rawls's The Law of Peoples." *Political Theory* 43 (4):473-499.

Krasnoff, Larry. 1999. "How Kantian is Constructivism?" *Kant-Studien* 90 (4):385-409.

Krishnamurthy, Meena. 2013. "Completing Rawls's Arguments for Equal Political Liberty and Its Fair Value: The Argument from Self-respect." *Canadian Journal of Philosophy* 43 (2):179-205.

Kolodny, Niko. 2010. "The Explanation of Amour-propre." *Philosophical Review* 119 (2):165-200.

Kuper, Andrew. 2000. "Rawlsian Global Justice: Beyond the Law of Peoples to a Cosmopolitan Law of Persons." *Political Theory* 28 (5):640-674.

Labukt, Ivar. 2009. "Rawls on the Practicability of Utilitarianism." *Politics, Philosophy and Economics* 8 (2):201-221.

Lane, Robert. E. 1982. "Government and Self-Esteem." *Political Theory* 10 (1):5-31.

Legge, James. 1879. *The Sacred Books of China: The Texts of Confucianism*. Oxford: Clarendon Press.

Levine, Andrew. 1974. "Rawls' Kantianism." *Social Theory and Practice* 3 (1):47-63.

Lyons, David. 1972. "Rawls Versus Utilitarianism." *Journal of Philosophy* 69 (18):535.

Macleod, Alistair M. 2006. "Rawls's Narrow Doctrine of Human Rights" in *Rawls's Law of Peoples: A Realistic Utopia?*, edited by Rex Martin & David Reidy, 134-149. Blackwell.

Maffettone, Pietro. 2011. "The Law of Peoples: Beyond Incoherence and Apology." *Journal of International Political Theory* 7 (2):190-211.

Marks, Jonathan. 2007. *Perfection and Disharmony in the Thought of Jean-Jacques Rousseau*.

Cambridge: Cambridge University Press.

Masters, Roger D. 1968. *The Political Philosophy of Rousseau*. Princeton: Princeton University Press.

Mason, H. E. 1976. "On the Kantian Interpretation of Rawls' Theory." *Midwest Studies in Philosophy* 1 (1):47-55.

McCloskey, H. J. 1957. "An Examination of Restricted Utilitarianism." *Philosophical Review* 66 (4):466-485.

Miller, D. 2008. "Political Philosophy for Earthlings." In *Political theory: Methods and Approaches*, edited by D. Leopold & M. Stears, 29–48. Oxford: Oxford University Press.

Miller, Richard W. 2012. "Rawls and Global Justice: A Dispute over a Kantian Legacy." *Philosophical Forum* 43 (3):297-309.

Moellendorf, Darrel. 2002. *Cosmopolitan Justice*. Westview Press.

Mohanty, Manoranjan. 2007. *Grass-roots Democracy in India and China: The Right to Participate*. New Delhi: Thousand Oaks, Calif.

Muchnik, Pablo. 2009. *Kant's Theory of Evil: An Essay on the Dangers of Self-Love and the Aprioricity of History*. Lanham: Lexington Books.

Narveson, Jan. 2010. "Cohen's Rescue." *Journal of Ethics* 14 (3-4):263-334.

Neuhouser, Frederick. 2008. *Rousseau's Theodicy of Self-Love: Evil, Rationality, and the Drive for Recognition*. Oxford University Press.

Neuhouser, Frederick. 2011. "Jean-Jacques Rousseau and the Origins of Autonomy." *Inquiry* 54 (5):478 - 493.

Neuhouser, Frederick. 2014. *Rousseau's Critique of Inequality: Reconstructing the Second*

Discourse.

Nicholas Tampio. 2007. "Rawls and the Kantian Ethos." *Polity, 39*(1): 79-102.

Nielsen, Kai. 1979. "Radical Egalitarian Justice." *Social Theory and Practice* 5 (2):209-226.

Nussbaum, Martha. 2002. "Women and the Law of Peoples." *Politics, Philosophy and Economics* 1 (3):283-306.

Nussbaum, Martha. 2006. *Hiding From Humanity: Disgust, Shame, and the Law*. Princeton University Press.

O'Hagan, Timothy. 1997. *Jean-jacques Rousseau and the Sources of the Self. Aldershot*, Hants, England: Avebury.

O'Hagan, Timothy. 1999. *Rousseau*. London: Routledge.

O'Neill, Onora. 1989. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge Univ. Press.

Qvortrup, Mads. 2003. *The Political Philosophy of Jean-Jacques Rousseau: The Impossibility of Reason*. Manchester: Manchester University Press**.**

Penny, Richard. 2013. "Incentives, Inequality and Self-Respect." *Res Publica* 19 (4):335-351.

Pogge, Thomas. 1981. "The Kantian Interpretation of Justice as Fairness." *Zeitschrift für Philosophische Forschung* 35 (1):47 - 65.

Pogge, Thomas. 2004. "Real World Justice." *Journal of Ethics* 9 (1-2):29-53.

Pogge, Thomas. 2007. *John Rawls: His Life and Theory of Justice*. Oxford University Press.

Pogge, Thomas. 2008. "Cohen to the Rescue!" *Ratio* 21: 454–475.

Pritchard, Michael S. 1977. "Rawls's Moral Psychology." *Southwestern Journal of Philosophy* 8 (1):59-72.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.

Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.

Rawls, John. 1999. *The Law of Peoples*. Cambridge, Mass.: Harvard University Press.

Rawls, John, 1999, *Collected Papers*, edited by Samuel Freeman. Cambridge, Mass.: Harvard University Press.

Rawls, John. 2000. *Lectures on the History of Moral Philosophy*, edited by Barbara Herman. Cambridge, Mass.: Harvard University Press.

Rawls, John. 2001. *Justice as Fairness*. Cambridge, Mass.: Harvard University Press.

Rawls, John. 2007. *Lectures on the History of Political Philosophy*, edited by Samuel Freeman. Cambridge: Belknap Press of Harvard University Press.

Reidy, David A. 2000. "Rawls's Wide View of Public Reason: Not Wide Enough." *Res Publica* 6 (1):49-72.

Reidy, David A. 2004. "Rawls on International Justice: A Defense." *Political Theory* 32 (3):291-319.

Reidy, David A. 2006. "Political Authority and Human Rights". in *Rawls's Law of Peoples: A Realistic Utopia?*, edited by Rex Martin & David Reidy, 169-188. Blackwell.

Reidy, David A. 2007. "A Just Global Economy: In Defense of Rawls." *Journal of Ethics* 11 (2):193-236.

Reidy, David A. 2009. "Rawls's Religion and Justice as Fairness." Manuscript.

Reidy, David A. forthcoming. "Moral Psychology, Stability and the Law of Peoples." Manuscript.

Roche, Timothy D. 1982. "Utilitarianism versus Rawls." *Social Theory and Practice* 8 (2):189-212.

Roosevelt, Grace G. 1941. *Reading Rousseau in the Nuclear Age*. Philadelphia: Temple University

Press.

Rousseau, Jean-Jacques. 1979. *Emile: Or, on Education*. New York: Basic Books.

Rousseau, Jean-Jacques. 1997. *The Social Contract*, in *The Social Contract and other Later Political Writings*, trans. Victor Gourevitch, 39-152. Cambridge: Cambridge University Press.

Rousseau, Jean-Jacques.1997. Discourses on Inequality (*Discourse on the Origin and Foundations of Inequality among Men*), in *The Discourses and Other Early Political Writings*, trans. Victor Gourevitch, 111-222. Cambridge University Press.

Rousseau, Jean-Jacques. 1997. *Discourse on Political Economy*, in *The Social Contract and Other Later Political Writings*, trans. Victor Gourevitch, 3-38. Cambridge University Press.

Rousseau, Jean-Jacques. 2001. *Rousseau, Judge of Jean-Jacques: Dialogues*, trans. Judith R. Bush, Christopher Kelly, and Roger D. Masters, in *The Collected Writings of Rousseau*, vol. 1. Hanover, NH: University Press of New England.

Rousseau, Jean-Jacques. 2001. *The Collective Writings of Jean-Jacques Rousseau*, trans. Judith R. Bush, Christopher Kelly, and Roger D. Masters. Hanover, NH: University Press of New England.

Sandel, Michael J. 2010. *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.

Shklar, Judith N. 1969. *Men and Citizens: A Study of Rousseau's Social Theory*. Cambridge: Cambridge University Press.

Smart, J. J. C. 1956. "Extreme and Restricted Utilitarianism." *Philosophical Quarterly* 6 (25):344-354.

Sinopoli, Richard C. 1994. “Association Freedom, Equality, and Rights Against the State”. *Political Research Quarterly*, 47(4):891-908.

Sivestrini, Gabriella. 2010, “Rousseau, Pufendorf and the Eighteenth-century Natural Law Tradition.” *History of European Ideas*, 36 (3): 280-301.

Strauss, Leo. 1965. Natural Right and History. Chicago: University of Chicago Press.

Sutton, Agneta. 1979. “The Kantian and the Consequentialist Elements in Rawls's Theory of Justice.” *Theoria* 45 (3):135-140.

Talisse, Robert B. 2000. “Rawls on Pluralism and Stability.” *Critical Review* 15 (1-2):173-194.

Taylor, Robert S. 2003. “Rawls’s Defense of the Priority of Liberty: A Kantian Reconstruction.” *Philosophy and Public Affairs* 31 (3):246–271.

Taylor, Robert S. 2011. *Reconstructing Rawls: The Kantian Foundations of Justice as Fairness*. Pennsylvania State University Press.

Thomas, Larry L. 1978. “Morality and Our Self-concept.” *Journal of Value Inquiry* 12 (4):258-268.

Thomas, Alan. 2011. “Cohen's Critique of Rawls: A Double Counting Objection.” *Mind* 120 (480):1099-1141.

Thompson, Norma. 2000. *Instilling Ethics*. Lanham, Md: Rowman & Littlefield.

Vaughan, Charles E. 1962. “Introduction” in *The Political Writings of Jean Jacques Rousseau* (1915), Vol. I: 16-7.

Viroli, Maurizio. 1988. *Jean-Jacques Rousseau and the “Well-Ordered Society”*. Cambridge: Cambridge University Press.

Weatherley, Robert. 2014. *Making China Strong: The Role of Nationalism in Chinese Thinking on Democracy and Human Rights*. Basingstoke: Palgrave Macmillan.

Wenar, Leif. 2001. “Contractualism and Global Economic Justice.” *Metaphilosophy* 32 (1-2):79-94.

Wenar, Leif. 2005. “The Unity of Rawls’s Work” in *The Legacy of John Rawls*, edited by Thom Brooks and Fabian Freyenhagen. Continuum, London.

Wenar, Leif. 2006. “Why Rawls is not a Cosmopolitan Egalitarian.” in *Rawls's Law of Peoples: A Realistic Utopia?*, edited by Rex Martin & David Reidy, 95-114. Blackwell.

Weithman, Paul J. 2010. *Why Political Liberalism?: On John Rawls's Political Turn*. Oxford University Press.

Wilfried Hinsch and Markus Stepanians. 2006. “Human Rights as Moral Claim Rights” in, *Rawls's Law of Peoples: A Realistic Utopia?*, edited by Rex Martin & David Reidy, 117-133. Blackwell.

Williams, Andrew (1998). “Incentives, Inequality, and Publicity.” *Philosophy and Public Affairs* 27 (3):225-247.

Yanal, Robert J. 1987. “Self-esteem.” *Noûs* 21 (3):363-379.

Ypi, Lea. 2012. “Facts, Principles and the Third man.” *Socialist Studies* 8: 196–215.

Zink, James. 2011. “Reconsidering the Role of Self-respect in Rawls’s A Theory of Justice.” *The Journal of Politics,* 73(2):331-344.

VITA

Xinghua Wang was born in a village of Changsha, Hunan Province of China, to the parents of Shoubin and Juyun. After finishing her high school in Changsha, she went north to pursue her bachelor degree at Wuhan University where she began to have a strong interest in philosophical thinking. She was one of the top students in Wuhan University — in the year in which she graduated she was awarded as the three good student. After she graduated in Wuhan, she was recommended to the department of philosophy at Renmin University of China where she began to have a keen interest in the problems of happiness and justice. She received her Master degree at Renmin University in May 2012. In the same year, her Master thesis won a university reward. She continued her study in the University of Tennessee, Knoxville where she was offered a teaching assistantship. During the years of her PhD, she attended two philosophy conferences in the United State and one in Romania. The paper presented at the conference in Romania was later revised and published. She did her doctoral thesis under the supervision of Dr. David Reidy. The thesis focuses on the significance of Rousseau's concept of amour-propre in Rawls.