

Congruence, Rationality and Teleology

Po Chung CHOW (周保松)

Department of Government and Public Administration, CUHK

p.c.chow@cuhk.edu.hk

Abstract

The problem of stability is fundamental to Rawls's political philosophy. In *Theory of Justice*, Rawls proposes a notion of congruence to resolve this problem. The central idea is that a conception of justice is stable if and only if the desire to act justly is congruent with the desire to realize our rational good. Only such congruence could ensure that rational members of a well-ordered society would have sufficient reason to give the first priority to the sense of justice.

This essay sets out to examine Rawls's argument for congruence. I shall first explicate the main ideas of congruence and show that its main ground lies in a Kantian interpretation of justice as fairness. I then argue that this interpretation has turned Rawls into a liberal perfectionist within a classical teleological framework. This position is, however, inconsistent with Rawls's desire-based conception of deliberative rationality. For this conception of rationality does not warrant that rational persons would necessarily accept a Kantian conception of the good. I shall conclude that it is this internal inconsistency which makes Rawls's argument for congruence fundamentally flawed and accounts for his philosophical turn to political liberalism.

Author's Note: Sections 3 & 4 are omitted from this abridged version. A full version is available upon request.

Congruence, Rationality and Teleology

Po-Chung Chow

The problem of stability is fundamental to Rawls's political philosophy. In *Theory of Justice*, Rawls proposes a notion of congruence to resolve this problem. The central idea is that a conception of justice is stable if and only if the desire to act justly is congruent with the desire to realize our rational good. Only such congruence could ensure that rational members of a well-ordered society would have sufficient reason to give the first priority to the sense of justice.

This chapter sets out to examine Rawls's argument for congruence. I shall first explicate the main ideas of congruence and then show that its main ground lies in a Kantian interpretation of justice as fairness. I argue that this interpretation has turned Rawls into a liberal perfectionist within a classical teleological framework. This position is, however, inconsistent with Rawls's desire-based conception of deliberative rationality. For this conception of rationality does not warrant that rational persons would necessarily accept a Kantian conception of the good. I shall conclude that it is this internal inconsistency which makes Rawls's argument for congruence fundamentally flawed and accounts for his philosophical turn to political liberalism.¹

This is a long chapter, not only in pages. The whole essay consists of eight sections,

¹ As Rawls explains, "to understand the nature and extent of the differences, one must see them as arising from trying to resolve a serious problem internal to justice as fairness, namely from the fact that the account of stability in Part III of *Theory* is not consistent with the view as a whole. I believe all differences are consequences of removing that inconsistency. Otherwise these lectures take the structure and content of *Theory* to remain substantially the same." (PL:xvii-xviii)

which are arranged as follows. Section 1 will introduce the main idea of congruence and shows its connection to moral stability. Since the need of congruence has been challenged by Barry as based on a misunderstanding of Ross's doctrine of purely conscientious act, Section 2 is set to compare Ross's doctrine with Rawls's theory of practical reason for action. I argue that deeply influenced by Foot, Rawls believes that the good of the sense of justice is required to justify the motivational priority of justice. In Sections 3 and 4, I start to examine Rawls's free-rider argument and social union argument for congruence. I shall argue that both arguments fail. The free-rider argument fails because an egoist would not act on moral reasons. The social union argument fails because Rawls's account of social union is insufficient to establish a shared final end for the priority of the sense of justice. What Rawls ultimately relies on is the Kantian interpretation argument, which is the focus of Section 5. Through careful exposition, I shall argue that this interpretation is a type of perfectionism grounded a conception of human nature as free and equal rational being. Moreover, my discussion about the second-order conception of personal autonomy in turn brings out an important distinction between neutral freedom and good freedom primarily suggested by Sidgwick in his commentary on Kant's moral philosophy. In Section 6, I suggest that the Kantian interpretation embodies both conceptions of freedom. However, these conceptions are internally incompatible with each other, and the pursuit of congruence requires Rawls to keep only the conception of good freedom. In Section 7, I shall show that the conception of good freedom has further revealed the Kantian interpretation as a classical teleological theory. The dichotomy between modern teleology and deontology obscures the teleological nature of Rawls's theory. Lastly, in Section 8, I maintain that there is an inconsistency between Rawls's teleology and his account of deliberative rationality, which results in the failure of congruence. My argument shall then shed light on Rawls's later development of political liberalism.

1 The Idea of Congruence

To begin with, I shall recapitulate the main ideas of congruence. Congruence refers to the convergence of two distinct standpoints, namely the standpoint of justice and the standpoint of goodness. The former is defined by principles chosen in the original position of equality by free and rational individuals under the reasonable constraints, while the latter is defined by successful execution of people's plans of life consistent with the criteria of deliberative rationality. As rational moral beings, we are guided by both standpoints in practical reasoning. The moral standpoint dictates us to do what justice requires while the rational standpoint moves us to realize our informed desire. They form the basis from which "institutions, actions and plans of life can be assessed." (TJ: 567/496-497 rev.) Both standpoints are legitimate, prescriptive and action-guiding.

Nevertheless, these standpoints do not always coincide with each other. An action which is right is not necessarily truly good for a rational agent. When the standpoints diverge from each other, the question of why it is rational for a person to act justly arises. This question has long been raised by Plato in *Republic*: "do you really want to convince us that right is in all circumstances better than wrong or not?"² According to Foot, if we cannot commend justice to people as a good, then "justice can no longer be recommended as a virtue."³ As a result, the motivational priority of justice will be unfounded. The existence of two standpoints and the possibility of conflict between them set the background for congruence.

To argue for the overridingness of the sense of justice, Rawls suggests that under

² Plato, *Republic*, ?

³ Philippa Foot, "Moral Beliefs" in *Virtues and Vices* (Oxford: Oxford University Press, 2002), p.125.

certain ideal condition, these two perspectives can converge. In that case, rational persons would have sufficient motive to abide by principles of justice. The basic question of congruence, according to Freeman, is that:

Is it rational in a well-ordered society of justice as fairness for persons to affirm individually, from the point of view deliberative rationality, the principles of justice they would rationally agree to when they take up the public perspective of justice?⁴

Rawls's answer is that in a well-ordered society regulated by justice as fairness, the sense of justice will be perceived as a regulative good by rational people. In such a society, Rawls believes that:

Not only does it generate its own supportive moral attitudes, but these attitudes are desirable from the standpoint of rational persons who have them when they assess their situation *independently from the constraints of justice*. (TJ:398-399/350 rev., emphasis added)

Thus congruence is a decisive factor in determining stability. It represents a harmony between moral life and good life. The demand of justice is not regarded as an external constraint imposed on a rational free agent. Instead, it is presented as a good constitutive to the agent's rational plan of life. Congruence expresses a distinctive ethical view that the practical reason for action must stem from one's conception of the good. To establish the priority of the sense of justice, we need to show how it can occupy a central place in an agent's subjective motivational set.⁵ This view has been seriously challenged by Barry. According to Barry, Rawls has in effect denied such a claim that recognizing something to

⁴ Freeman, "Congruence and the Good of Justice", p.285

⁵ This is quite close to what Williams called the Sub-Humean internalist view. *Moral Luck* (Cambridge: Cambridge University Press, 1981), pp.101-13.

be right is sufficient to motivate right action.⁶ An agent would not do what is right simply out of a sense of duty. Barry suggests that Rawls's view stems from his rejection of the doctrine of pure conscientious act held by Ross, according to which "the highest moral motive is the desire to do what is right, and just simply because it is right and just, no other description being appropriate." (TJ:477/418 rev.) Having rejected this peculiar doctrine about moral motive, Barry comments:

Rawls commits himself in Chapter 9 of *A Theory of Justice* to the ancient doctrine that no act can be regarded as rational unless it is for the good of the agent to perform it. Thus, the problem is one of "congruence argument between justice and goodness."⁷

Barry argues that Rawls makes an apparent mistake because the desire to act morally for the sake of justice is a widespread disposition in society. People simply fulfill their duty out of their sense of duty and not for some independently defined good. Furthermore, the idea of congruence is unnecessary and absurd. It is unnecessary because Rawls need not rely on it to solve the problem of stability. In Chapter 8 of *Theory of Justice*, Rawls has already shown that following the three laws of moral psychology people will develop an effective sense of justice to do what justice requires. Barry believes that this widely shared moral disposition is strong enough to maintain a stable society without recourse to the demanding congruence argument. Rawls's claim is absurd because it will dissolve an important distinction between the right and the good in our practical reasoning, to wit, we are capable of doing something right while believing that what morally forbids is good for us. Barry offers an example to illustrate this point:

⁶ Barry, "John Rawls and the Search for Stability," p.884.

⁷ Barry, "John Rawls and the Search for Stability," p.885.

Suppose that I form the view that it would contribute to my good to take a trip around the world, and that I then find that this would cost more than my resources permit...Instead of simply concluding that I cannot justly take the trip (while continuing to believe that taking it would be for my good), I am told by Rawls that I must somehow persuade myself that it would not be for my good at all. *For only that thought can motivate me to refrain from taking the trip unjustly if the opportunity should arise.* This is the absurdity into which Rawls is led by his rejection of “the doctrine of the purely conscientious act.”⁸

Barry thus concludes that the whole idea of congruence is grounded on a mistake. It can be set aside without causing any damage to Rawls’s theory as a whole. It follows that it is also a mistake for Rawls to turn to political liberalism.⁹ I think Barry’s criticism is unfair to Rawls. Through a careful analysis of Ross’s doctrine and Rawls’s view of practical reason for action in the next section, I shall show that there is indeed a need for congruence.

2 The Need for Congruence

The doctrine of purely conscientious act is concerned with the nature and source of motivation of moral obligation. Ross claims that the goodness of morally good actions must arise from a certain kind of motive which is connected with a certain type or types of character. This kind of motivation is our sense of duty. It is stipulated as something distinct from, and superior to, other desires of any kind.¹⁰ Only so will the necessity of

⁸ Barry, “John Rawls and the Search for Stability,” p.889, my emphasis.

⁹ In the last sentence of his long article, Barry concludes that “Rawls’s sweeping recantation is uncalled-for, and that the failure of *Political Liberalism* does not discredit *A Theory of Justice*. I believe that, as time goes on, *A Theory of Justice* will stand out with increasing clarity as by far the most significant contribution to political philosophy produced in this century. Only one thing threatens to obscure that achievement: the publication of *Political Liberalism*.” Barry, “John Rawls and the Search for Stability,” p.915.

¹⁰ At this point, Ross disagrees with Kant that apart from the sense of duty, there are other desires resulting from our nature as rational beings. He thus says that “we can agree with him (Kant) in thinking that the sense of duty is the highest motive, without following him in putting all other motives on the same dead level.” *Foundations of Ethics* (Oxford: the Clarendon Press, 1939), p.206.

moral obligation and its supreme worth of conscientious actions be warranted. Where does this moral motive stem from? From our pure practical reason! The recognition of an act as one's duty can by itself motivate us to act:

There is no more mystery in the fact that the thought of an act as one's duty should arouse an impulse to do it, than in the fact that the thought of an act as pleasant, or as leading to pleasure, should arouse an impulse to do it.¹¹

Ross contends that this desire to do our duty need not presuppose Kant's metaphysical view about human beings' phenomenal and noumenal nature. Rather, it simply springs from our possession of reason. As he puts, "it is only natural that there should arise a desire, itself springing from our rational apprehension of principles of duty, not to be the slave of low desires but to regulate our life by these principles."¹² A purely conscientious act proceeds from "*a desire for a specifically distinct object*, not for the attainment of pleasure nor even for the conferring of it on others, but just for the doing of our duty."¹³ Ross stresses that the word "purely" specifies the distinct nature of singleness of this desire.¹⁴

This is essentially a Kantian account of moral motive. A morally good action must act from a sense of duty, which is detached from other inclinations or a person's happiness. "An action from duty is to put aside entirely the influence of inclination and with it every object of the will."¹⁵ There is no relation between the moral motive and the good. So, a rational agent should comply with the moral law even if it infringes upon all his inclinations and interests. Thus the priority of the sense of duty is purely based on our

¹¹ W. D. Ross, *The Right and the Good* (Indianapolis: Hackett Publishing Company, 1930), pp.157-58.

¹² Ross, *Foundations of Ethics*, p.206.

¹³ *The Right and the Good*, p.158, emphasis added.

¹⁴ *Foundations of Ethics*, p.207.

¹⁵ Kant, *Groundwork of the Metaphysics of Morals* collected in *Practical Philosophy*, tran. & edited by Mary J. Gregor (Cambridge: Cambridge University Press, 1996), p.55.

nature of practical reason. It should be noted that it is this particular Kantian position that Rawls describes as the doctrine of purely conscientious act, but not Ross's more general doctrine of moral intuitionism.¹⁶ He raises this issue not because he has a primary interest in Ross's intuitionism, but because he wants to make a comparison of this doctrine with his own view of moral motivation so that the distinctive feature of congruence and its indispensability to justice as fairness can be seen.¹⁷

What is wrong with Ross's account of moral motive then? Rawls offers the following explanation:

It would seem then, that the doctrine of the purely conscientious act is *irrational*...Ross holds that the sense of right is a desire for a distinct (and unanalyzable) object, since a specific (and unanalyzable) property characterizes actions that are our duty. The other morally worthy desires, while indeed desires for things necessarily connected with what is right, are not desires for the right as such. But on this interpretation *the sense of right lacks any apparent reason*; it resembles a preference for tea rather than coffee. Although such a preference might exist, to make it *regulative* of the basic structure of society is utterly *capricious*. (TJ:477-78/418 rev., emphasis added)

At first sight it is unclear why this doctrine is irrational and capricious. If we agreed with Barry, there would be nothing wrong to hold that recognition of something as just can give rise to a sense of duty. Mendus tries to defend Rawls by saying that what Rawls specifically objects to is the concept of goodness (or rightness, or justice) as simple and unanalysable.¹⁸ It is a rather peculiar account of moral motivation of rational intuitionism. Barry's charge is thus harmless because Rawls could agree with him on the relation

¹⁶ Rawls has clearly indicated his reference to this view in his discussion. (TJ: 477/418 rev., footnote 15)

¹⁷ Rawls has never used the term "intuitionism" to describe Ross's theory. He refers intuitionism to such a doctrine that "there is an irreducible family of first principles which have to be weighted against one another by asking ourselves which balance, in our considered judgment, is the most just." (TJ:34/30 rev.)

¹⁸ Mendus, "The Importance of Love in Rawls's *Theory of Justice*," *British Journal of Political Science*, 29, (1999), p.62.

between motivation and practical reason in general.¹⁹ Nevertheless, this defense is not of much help because it cannot explain why Rawls would uphold the demanding requirement of congruence. Barry, for instance, may argue that what Rawls should do is simply to accept a Scanlonian account of moral motive, and abandon the project of congruence. Besides, we should note that Rawls's complaint is that Ross's doctrine act is *irrational*. He acknowledges that such motive may exist. What he argues against is that it is taken as a regulative motive of the basic structure. So, more attention should be paid to the problem of rationality.

To judge whether an act is rational, we need to know Rawls's theory of rationality. Rawls holds a conception of deliberative rationality, which stipulates that an act is rational when it can best achieve what an agent wants most after informed deliberation. The agent's rational desire ultimately determines his practical reason for action. This is what Parfit called a *desired-based* theory.²⁰ One of the distinctive features of this theory is that rationality itself cannot decide what an agent should most want. Therefore, "knowing that people are rational, we do not know the ends they will pursue, only that they will pursue them intelligently." (PL:49, footnote 1) Regardless of this, Rawls makes a further assumption that we all live according to a plan of life. One's rational plan of life determines his good. A plan is rational if, and only if:

(1) it is one of the plans that is consistent with the principles of rational choice when these are applied to all the relevant features of his situation, and (2) it is that plan among those meeting this condition which would be chosen by him with full deliberative rationality, that is, with full awareness of the relevant facts and after a careful consideration of the consequences. (TJ: 408/358-359, rev.)

¹⁹ Freeman deploys the similar strategy to respond to Barry. Freeman, "Congruence and the Good of Justice," p.282.

²⁰ Derek Parfit, "What We Could Rationally Will," *The Tanner Lectures on Human Value*, (?), p.342

So, an act is rational if and only if it is the most effective way to realize what the agent most wants, namely, the plan of life that he will adopt with full deliberative rationality. It implies that *in most cases* a rational agent will only have a motive to act on principles that is beneficial to his plan of life.²¹ When the question “why should I be just” is raised, the answer hinges considerably on whether acting justly can be conceived as a good to the rational agent.²² “The desire to act justly is not, then, a form of blind obedience to arbitrary principles unrelated to rational aims.” (TJ:476/417 rev.) To the contrary, Rawls argues:

A theory should present a description of an ideally just state of affairs, a conception of a well-ordered society such that the aspiration to realize this state of affairs, and to maintain it in being, *answers to our good and is continuous with our natural sentiments.* (TJ:477/417 rev., emphasis added)

We can now understand why Rawls thinks that the motive assumption of purely conscientious act is irrational. For the doctrine fails to provide any further justification to rational individuals that they have a duty to give absolute priority to the moral sentiment over other desires. It simply supposes that the highest moral motive is the desire to do what is right and just simply because it is right and just. The desire to do our duty is therefore completely detached from the desire to realize our plan of life. When a person asks himself why it is rational for him to fulfill his duty even if there is a strong conflict between moral command and his conception of the good, Ross’s reply is that “the truest answer I can find is that I do it because, then at least, I desire to do my duty more than I

²¹ I emphasize “in most cases” because Rawls says that it is still rational for a person to lose his life for the sake of justice if, all things considered, this is what he most wants to do. (TJ:573/502) Thus Rawls is not an egoist. If the good is broadly defined as the satisfaction of what one most wants, the discrepancy will disappear.

²² As Rawls himself acknowledges, his view is greatly influenced by Foot’s seminal article “Moral Belief” in which Foot argues that we must take the question of “why should I be just” seriously. Foot, “Moral Belief,” pp.125-30.

desire anything else.”²³ For Rawls, this answer is question-begging. It lacks any normative force to establish its motivational authority.

Justice is a virtue concerning what we owe to other people. It may require us to sacrifice our great interests for the sake of other’s rights. Unlike other actions that would bring pain, boredom and loneliness, the desire to act justly could hardly give sufficient reasons for action by itself. As Foot remarks, “ ‘it is unjust’ gives a reason only if the nature of justice can be shown to be such that it is necessarily connected with what a man wants.”²⁴ Foot believes that there are two types of desires for actions. The first type can directly motivate us in a certain way because it can satisfy some of our basic wants that are ultimately related to our well-being. In this case, no further justification is required. Avoiding pain is such an example.²⁵ On the contrary, the second type needs *further* reason to trigger our action. The sense of justice belongs to this type. Rawls shares with Foot:

The desire to act justly is not a final desire like that to avoid pain, misery, or apathy, or the desire to fulfill the inclusive interest. The theory of justice supplies *other descriptions* of what the sense of justice is a desire for. (TJ:569/499 rev., emphasis added)

For Rawls, the “other descriptions” are mainly related to how an individual, consistent with the desire-based deliberative rationality, can confirm the sense of justice as a regulative desire of his plan of life. It is for this reason Rawls finds the doctrine of purely conscientious act irrational, and congruence necessary to moral stability. Rawls not only

²³ Ross, *Foundations of Ethics*, p.206.

²⁴ Foot, “Moral Belief,” p.127.

²⁵ This point is famously illustrated by Hume in the following paragraph. “Ask a man *why he uses exercise*; he will answer, *because he desires to keep his health*. If you then enquire, *why he desires health*, he will readily reply, *because sickness is painful*. If you push your enquires farther, and desire a reason *why he hates pain*, it is impossible he can ever give any. This is an ultimate end, and is never referred to any other object.” Hume, *Enquires*, Appendix I, V.

rejects Ross's specific account of moral motive, but also disapprove of a general position that the sense of justice detached from our good can effectively motivate rational individuals to act.

Barry may argue that the sense of justice does not belong to the second type of desire. On the contrary, as a matter of simple fact, "people can do their duty out of a sense of duty and not in order to achieve some independently definable end."²⁶ Based on this conviction, Barry could still deem Rawls's account of reason for action flawed. However, Barry misses Rawls's point. Rawls does not deny that a rational individual may act for the sake of justice. He recognizes this possibility before bringing out the real problem of congruence:

Supposing that someone has an effective sense of justice, he will then have a regulative desire to comply with the corresponding principles...Therefore in this form the question is trivial: being the sorts of persons they are, the members of a well-ordered society desire more than anything to act justly and fulfilling this desire is part of their good. (TJ:569/498 rev.)

If the majority of citizens develop this effective desire to act upon principles of justice for its own sake, Rawls would have every reason to be happy because this would bring the most stable society that one can hope for. Nevertheless, Rawls holds that a stable well-ordered society should not rely too much on this possibility. He is not as optimistic as Barry that people would have a general and regulative sense of justice for its own sake. There is a real disagreement between them about moral psychology and the source of normativity.

We may also think about this question from the perspective of circumstance of justice.

²⁶ Barry, "John Rawls and the Search for Stability," p. 884.

Rawls presumes that society is a cooperative venture for mutual advantage. All members are supposed to have a fundamental interest in advancing their conceptions of the good. Their primary motive to participate in cooperation is to realize their good. This explains why they have conflicting claims about distributive justice. Rational individuals are not expected to act from a pure sense of moral duty. To what extent they can develop an effective moral motive largely depends upon whether the principles of justice concerned can effectively promote their good. Therefore, the real problem of congruence is about “what happens if we imagine someone to give weight to his sense of justice only to the extent that it satisfies other descriptions which connect it with reasons specified by the thin theory of the good.” (TJ:569/499 rev.)

Given what I have argued in this section, Rawls’s project is not as unnecessary and absurd as Barry perceives. On the contrary, the quest for convergence is a necessary step to achieve moral stability. Before we go to assess Rawls’s substantive argument for congruence, an ambiguity concerning the definition of the good must be settled. Rawls repeatedly reminds us that when he talks about the good of the sense of justice, the good must be understood in a thin sense. This is a puzzling reminder. We know that the main purpose of the thin theory of the good is to define primary goods in the original position. Primary goods are those things that rational individuals, whatever they are, desire as prerequisites for carrying out their plans of life. These things include rights and liberties, income and wealth, opportunity, and self-respect. Nevertheless, the sense of justice does not belong to this list. Moreover, the thin theory is mainly used in the original position. After the veil of ignorance is lifted, people will deliberate according to the full theory of the good in connection with deliberative rationality and their plans of life. Now since congruence takes place in the second stage where people have full knowledge of their

situation with no restriction on information, it is unreasonable to suppose that people would still be guided by the thin theory of the good in judging the goodness of the sense of justice. Barry, for example, has put the following question to Rawls:

Given that the problem, as Rawls conceives it, is one of offering a reason for real people to “affirm their sense of justice,” it is surely correct to specify that they should have full information. But then why should they throw away the information about their own distinctive conceptions of the good (their “thick” conception) and restrict the question to one the answer to which is going to be the same for everyone?²⁷

This is a good observation, and Rawls’s position is ambiguous. For example, he explicitly states that “when we come to the explanation of the social values and the stability of a conception of justice, *a wider interpretation of the good* is required.” (TJ:398/350 rev., emphasis added) He also says that the real problem of congruence is whether a rational person would adopt the standpoint of justice as a regulative good viewed “in the light of the thin theory *with no restrictions on information.*” (TJ:567/497 rev., emphasis added) If rational individuals have already known their conceptions of the good, it is implausible to expect them not to use the full theory of the good to assess their situation.

In my view, the only plausible explanation is that even though the knowledge of conceptions of the good is available to people, the thin theory should still be used to justify that it is rational for *every* person to desire the sense of justice as regulative good. It is thin in a sense that it is good for all rational plans of life. Therefore, “the plan of life which does this is his best reply to the similar plans of his associates; and *being rational for anyone, it is rational for all.*” (TJ:567/497 rev., emphasis) The real challenge is, given

²⁷ Barry, “John Rawls and the Search for Stability,” pp.885-86.

the desire-based rationality and the fact of pluralism, whether Rawls can offer substantive arguments for this goal. This is what we are going to examine in the rest of this chapter.

5. The Kantian Interpretation Argument

The central idea of the Kantian interpretation of justice as fairness is that rational persons are said to have a regulative desire to express their nature as free and equal rational being. The idea is primarily derived from Kant's conception of moral person and the notion of autonomy. Rawls holds that justice as fairness can be interpreted as the most adequate expression of this Kantian ideal. It provides the strongest reason for rational agents to uphold justice because acting justly is itself a regulative good judging from the rational point of view. In the following discussion, I will first explicate this interpretation and then argue that it has presupposed a form of perfectionism.

It would be instructive to begin with an examination of how the right and the good are congruent with each other under the Kantian interpretation of human nature. According to Rawls:

Acting justly is something we want to do as free and equal rational beings. The desire to act justly and the desire to express our nature as free moral persons turn out to specify what is practically speaking the same desire. When someone has true beliefs and a correct understanding of the theory of justice, these two desires move him in the same way. They are both dispositions to act from precisely the same principles: namely, those that would be chosen in the original position.
(TJ:572/501 rev.)

Having explained the possible convergence of two desires, Rawls continues to argue for its regulative priority:

The desire to express our nature as a free and equal rational being can be fulfilled only by acting on the principles of right and justice as having first priority. This is a consequence of the condition of finality: since these principles are regulative, the desire to act upon them is satisfied only to the extent that it is likewise regulative with respect to other desires. It is acting from this precedence that expresses our freedom from contingency and happenstance. Therefore, *in order to realize our nature we have no alternative but to plan to preserve our sense of justice as governing our other aims*. This sentiment cannot be fulfilled if it is compromised and balanced against other ends as but one desire among the rest. (TJ:574/503 rev., emphasis added)

Rawls's argument can be formulated as having the following thirteen steps.

- (1) According to the Kantian account of human nature, we are essentially free and equal rational beings.
- (2) Rational beings have a fundamental desire to express their nature.
- (3) The realization of one's nature is a supreme good for a rational person.
- (4) The necessary and sufficient condition of realizing one's nature is to act on principles of justice which are chosen under conditions that fully represent his nature as free and equal rational being.²⁸ This is because "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:253/222 rev.)
- (5) The original position is designed to fulfil the task in (4). As a device of representation, it specifies conditions in which principles of justice are chosen by rational free persons in an initial situation of equality without being affected by their social position or natural endowments.
- (6) Since Rawls's principles of justice would be chosen by free and equal parties in the original position, acting from these principles is therefore the most effectively way to express their nature.

²⁸ Rawls uses "to express one's nature" and "to realize one's nature" interchangeably to refer to the same meaning.

(7) The desire to apply and to act from the principles of justice is called a sense of justice. (TJ:567/496-497 rev.)

(8) Taking the above premises together, it can be concluded that “acting justly is something we want to do as free and equal rational beings. The desire to act justly and the desire to express our nature as free moral persons turn out to specify what is practically speaking the same desire.” (TJ: 572/501 rev.)²⁹

(9) Thus it is rational for an individual to affirm their sense of justice because the realization of one’s nature is a supreme good in virtue of (3).³⁰

(10) Furthermore, the condition of finality requires that the principles of justice chosen in the original position must be regulative and overriding. The primacy of justice is assured by this formal condition.

(11) Since the principles of justice are regulative, the desire to act upon them must also be regulative. If the sense of justice is compromised and balanced against other ends as but one desire among the rest, we fail to fully realize our nature. For “how far we succeed in expressing our nature depends upon how consistently we act from our sense of justice as *finally regulative*.” (TJ:575/503 rev., emphasis added)

(12) Therefore, it is always rational for individuals in a Rawlsian well-ordered society to give first priority to the sense of justice as a result of congruence.

(13) Finally, congruence is verified and justice as fairness is shown to be the most stable conception of justice.

²⁹ Similarly, after rejecting the doctrine of the purely conscientious act as irrational, Rawls asserts that “for one who understands and accepts the contract doctrine, the sentiment of justice is not a different desire from that to act on principles that rational individuals would consent to in an initial situation which gives everyone equal consideration as a moral person. Nor is it different from wanting to act in accordance with principles that express men’s nature as free and equal rational beings.” (TJ:478/418 rev.)

³⁰ Freeman holds that this is a result of the Aristotelian Principle. Nevertheless, according to Rawls, our good is determined by the satisfaction of our rational desires. Since the expression of our nature is a rational desire, it is not necessary to appeal to the Aristotelian Principle to affirm its value. Freeman, “Congruence and the Good of Justice,” pp.293-94.

We can note that the whole argument is grounded on a particular interpretation of human nature. If the argument holds true, justice as fairness will be as stable as one can hope for. There will be no disharmony between the right and the good. A well-ordered society regulated by justice as fairness is not only reasonable from the impartial perspective of the original position of equality, but also desirable from the first-person rational point of view. If one is rational enough, he will recognize that being a just person is exactly what he most desires to be. The congruence argument has a deep appeal because it offers an answer to the question of why a rational person would have sufficient reason to act morally.

This Kantian conception of human nature seems to point to a perfectionist position. According to Rawls, perfectionism is a type of ideal-regarding principle which directs society to arrange institutions and to define the duties and obligations of individuals so as to realize an ideal independently specified. (TJ:325-26/285-286 rev.)³¹ Thomas Hurka, a prominent perfectionist philosopher, also defines perfectionism as such a doctrine that “the good consists at bottom in developing one’s ‘nature’, or realizing a ‘true self’: certain properties are central to one’s identity, and one’s good consists in developing these properties to a high degree.”³² Furthermore, perfectionism is based on an objective theory of the good. The realization of good in a life makes that life better independently of how much they are wanted or enjoyed.

³¹ Rawls holds that there are two variants of perfectionism. In the first, it is the sole principle of a teleological theory which aims to maximize the achievement of human excellence; in the second, the principle of perfection is only one standard among several in an intuitionist theory. (TJ:325/285-286 rev.) This indicates that maximization is not a defining feature of perfectionism. What is crucial is the role of human excellence played in a theory. I will elaborate this point in more detail later.

³² Thomas Hurka, “Perfectionism” in E. Craig ed., *Routledge Encyclopedia of Philosophy*, (London: Routledge, 1998), retrieved from <http://www.rep.routledge.com/article/L070SECT5>.

The Kantian interpretation of justice as fairness seems to meet the requirement of perfectionism perfectly. It prescribes an objective conception of human nature as free and equal rational being; it holds that a justifiable conception of justice must be able to express our nature as fully as possible; it even stipulates that acting in accordance with justice is a constitutive good with the highest priority. At some point, Rawls accepts this characterization of his theory. For instance, he remarks that “a certain ideal is embedded in the principles of justice, and the fulfilment of desires incompatible with these principles has no value at all. Moreover we are to encourage certain traits of character, especially a sense of justice.” (TJ:326-27/287 rev.) Despite this, Rawls insists that “perfectionism is denied as a political principle.” (TJ:329/289 rev.) Justice as fairness at most occupies “an *intermediate position* between perfectionism and utilitarianism.” (TJ:327/287 rev., emphasis added)

Rawls offers two arguments to defend his position. First, he stresses that his ideal conception of moral person does not invoke “a prior standard of human excellence.” (TJ:327/287 rev.) This explanation is dubious. If justice as fairness is founded on a prior conception of human nature, it will entail a prior standard of excellence. The standard is to realize our nature as free and equal rational being as fully as possible. It determines how the original position would be designed and what kind of principles would be chosen; rational citizens are also expected to have a higher-order interest to develop their moral capacity for a sense of justice and for a conception of the good. “For this sentiment reveals what the person is.” (TJ:575/503 rev.) In this regard Rawls is undeniably a liberal perfectionist.

Rawls’s second argument is that any principles of perfection would be rejected by contractors in the original position because it fails to provide a firm basis for the equal

liberties in a pluralistic society. Although contractors are not cognizant of their particular moral and religious interests, they are aware that they will devote to different conceptions of the good in a well-ordered society. Any perfectionist principle will be incompatible with equal basic liberties for all. So, they would not “risk their freedom by authorizing a standard of value to define what is to be maximized by a teleological principle of justice.” (TJ:328/288-289 rev.)

The validity of this argument is based on an assumption that the principle of equal liberty itself does not depend upon any perfectionist ideal. This assumption is disputable though. It should be reminded that the reason for the contractors to give top priority to liberty is mainly that they regard themselves as free and equal rational beings. For Rawls, the principle of equal liberty is the only alternative compatible with our nature, and “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:253/222 rev., emphasis added) The original position serves as a mediating idea to represent our nature as free and equal. Our differences in social class and natural endowment are excluded from the original position because they would vitiate our nature as autonomous equal beings. To act on principles derived from these factors is to act heteronomously. (TJ:252/222 rev.) Similarly, the fact that we affirm a particular ideal does not give us a good reason to expect others to accept a conception of justice in that ideal’s favour because it is incompatible with our capacity to form, revise, and pursue a conception of the good. This explains why the contractors are to be ignorant of their philosophical and religious worldviews. In setting up the original position, the Kantian ideal has already been incorporated in its description. It plays a *decisive determining role* in deriving Rawls’s principles of justice. In this sense what the original position represents is not a *neutral*

point of view. Rather, it embodies a distinctive liberal conception of person. “The parties conceive of themselves as free persons who can revise and alter their final ends and who give priority to preserving their liberty in this respect.” (TJ:475, rev.)³³ My claim that Rawls is a liberal perfectionist thereby sustains.

That being said, it may be argued that Rawls’s perfectionism is a second-order ideal that can accommodate a variety of *first-order* substantive conceptions of the good. The Kantian ideal encourages people to develop their capacity as autonomous rational agency through choosing their ways of life. But it will not privilege any particular conceptions of the good at the ground-floor level. As Barry puts, “anything could be regarded as good (in a second-order way) so long as the person who conceived it as good (in a first-order way) had arrived at this conception in a way that satisfied the requirements of autonomy.”³⁴

This defence, if valid, seems to have the advantages of two worlds. On the one hand, it can admit that liberalism is grounded on a particular interpretation of human nature; on the other hand, it can avoid direct competition with other conceptions of the good by positing itself in a higher-order position. Many liberals hold this position. Kymlicka, for example, believes that this is the moral foundation of Rawls’s theory. Moreover, the value of personal autonomy does not lead to perfectionism. In contrast, state neutrality is required to respect people’s self-determination.³⁵ I am not convinced by this account. Nor do I believe that this is a proper interpretation of Rawls. In the rest of this section, I will

³³ The most fundamental change of revised edition of *Theory of Justice* is to re-define the parties in the original position to have a common higher interest in developing their two moral powers -- their capacity for a sense of justice and their capacity for a conception of the good -- in order to secure the priority of basic liberty. This indeed confirms my claim that the derivation of principle of liberty depends upon a perfectionist account of human nature and the corresponding human interest. (TJ: xii, rev.)

³⁴ Barry, *Justice as Impartiality*, p.129.

³⁵ Kymlicka, *Contemporary Political Philosophy* (Oxford: Clarendon Press, 1990), p.207.

argue that even Kymlicka himself has adopted a theory of liberal perfectionism.

According to Kymlicka, the starting point of liberalism is that we all have an essential interest in leading a good life. However, leading a good life differs from leading the life we currently believe to be good. For our belief about value could be mistaken. We may misunderstand our real interest or misjudge the value of a particular activity. It follows that we should be able to stand back from our existing ends and deliberate whether our plan of life is really worth pursuing. This does not mean that one who believes that he is in a better position to know what is good can impose his view on another person because “no life goes better by being led from the outside according to values the person does not endorse.”³⁶ This is what Kymlicka calls “endorsement constraint.” No matter how good a way of life may be from a third-person perspective, it cannot make a person’s life better if it is not accepted by that person from inside, according to his beliefs about value. In Kymlicka’s view, the endorsement constraint is applicable to most valuable forms of human activity.³⁷ As a consequence, leading a good life requires two pre-conditions:

One is that we lead our life from the inside, in accordance with our beliefs about what gives values to life; the other is that we be free to question those beliefs, to examine them in the light of whatever information, examples and arguments our culture can provide.³⁸

³⁶ Kymlicka, *Contemporary Political Philosophy*, p.203.

³⁷ Kymlicka, however, admits that sometimes short-term state intervention is justifiable if we accept that “one way to get people to pursue something for the right reasons is to get them pursue it for the wrong reasons, and hope that they will then see its true value.” Therefore, “the endorsement constraint argument, by itself, cannot rule out all forms of state perfectionism.” Once this qualification is granted, the liberal objection to perfectionism is no longer as strong as it primarily claims. *Contemporary Political Philosophy*, p.233.

³⁸ Kymlicka, *Liberalism, Community and Culture* (Oxford: Clarendon Press, 1990), p.13.

These conditions justify the value of self-determination. Kymlicka then claims that this account manifests Rawls's conception of free person. A free person is characterized as capable of forming, revising, and rationally pursuing a conception of the good.³⁹ In a nutshell, free persons "think of themselves not as inevitably tied to the pursuit of the particular final ends they have at any given time, but rather as capable of revising and challenging these ends on reasonable and rational grounds."⁴⁰ Rawls further stipulates that free persons are moved by a higher-order interest to exercise this distinctive power of self-determination. It is higher-order in a sense that it is supremely regulative and effective. It governs our deliberation and conduct whenever circumstances are relevant to its fulfilment.⁴¹ Why should we have such regulative desire to preserve this capacity? Kymlicka argues that Rawls's answer must be that we have an essential interest in leading a good life.⁴² Nevertheless, the commitment to self-determination does not lead to perfectionism. On the contrary:

Rawls argues that this account of self-determination should lead us to endorse a 'neutral state'—i.e. a state which does not justify its actions on the basis of the intrinsic superiority or inferiority of conceptions of the good life, and which does not deliberately attempt to influence people's judgments of the value of these different conceptions.⁴³

The connection between self-determination and neutrality is thereby established. Kymlicka calls this conception of neutrality *justificatory neutrality*.⁴⁴ This is however a

³⁹ Another feature is our capacity for a sense of justice, the capacity to understand, to apply, and to act from the public conception of justice. (PL:19)

⁴⁰ Rawls, *Collected Papers*, p.309.

⁴¹ Rawls, *Collected Papers*, p.312.

⁴² Kymlicka, *Liberalism, Community and Culture*, p.12.

⁴³ Kymlicka, *Contemporary Political Philosophy*, p.205.

⁴⁴ Another conception of neutrality is called "consequential neutrality" which requires that the state should seek to help or hinder different life-plans to an equal degree. Kymlicka, "Liberal Individualism and Liberal Neutrality" in *Communitarianism and Individualism* ed. Shlomo Avineri and Avner de-Shalit (New York:

mistake. By definition, a conception of the good normally consists of a more or less systematic account of what is valuable in human life. It provides a framework through which we can rank our preferences and give meaning to our life. It guides our action. According to Rawls, Kant and Mill's liberalism are examples of comprehensive conception of the good because they appeal to autonomy and individuality respectively to inform our thought and conduct as a whole. (PL: 37, 78)

In view of this, it is fair to say that Kymlicka's self-determination-based liberalism has presupposed a conception of the good. For instance, he has repeatedly borrowed support from Mill to vindicate the importance of autonomy in leading a good life.⁴⁵ He also objects to Rawls's political liberalism by arguing that the value of autonomy should not only be limited to the political sphere. Rather, it should govern human thought and action generally.⁴⁶ Kymlicka's message is clear: forms of life are truly valuable for us only if we perceive them as ones we endorse, or would endorse in a reflective and critical manner. To lead a good life, we must regard ourselves as autonomous beings who can freely and rationally question our beliefs. A liberal should deem exercising our capacity for self-determination a regulative interest, and respect for independence and individuality. Although autonomy allows a wide range of choice of different substantive views of the good life, it shapes our life in a fundamental way.

This liberal ideal is an object of reasonable disagreement though. It is not even widely shared in Western democratic societies. Kymlicka admits that there are many existing non-liberal minority groups who do not give priority to self-determination over

Oxford University Press, 1992), p.166.

⁴⁵ Kymlicka, *Liberalism, Community and Culture*, pp.9-19.

⁴⁶ Kymlicka, *Multicultural Citizenship* (Oxford: Oxford University Press, 1995), p.160.

their religious belief and cultural practice. Later Rawls is also well aware that many people living in a democratic society may not value autonomy at all:

Citizens may have, and normally do have at any given time, affections, devotions, and loyalties that they believe they would not, and indeed could and should not, stand apart from and objectively evaluate from the standpoint of their purely rational good. They may regard it as simply unthinkable to view themselves apart from certain religious, philosophical and moral convictions, or from certain enduring attachments and loyalties.⁴⁷

This indicates that in Rawls's mind, self-determination-based liberalism embodies a perfectionist conception of the good. Kymlicka is thus wrong to claim that self-determination requires state neutrality. Kymlicka counters that even though a liberal society encourages rational assessment and revisions of one's ends, it does not compel people to lead a particular form of life. Hence, "even if this view of autonomy conflicts with a religious minority's self-understanding, there is no cost to accepting it for political purpose."⁴⁸ Nevertheless, this is merely Kymlicka's own wishful thinking. The court case *Winsonsin vs. Yoder* which Kymlicka cites is a good example. The Amish community requests the government to allow them to withdraw their children from school before the age of 16 in accordance with its religious doctrine. Kymlicka argues that this kind of internal restriction is unacceptable because it violates children's individual autonomy. But from the Amish point of view, they are paying a heavy cost for living in a liberal society. Liberal autonomy is not as neutral as Kymlicka claims. As Barry rightly points out, this second-order conception of the good as autonomy actually requires that "only those conceptions that have the right origins—those that have come about in ways that meet the

⁴⁷ Rawls, "Justice as Fairness: Political not Metaphysical," p.214.

⁴⁸ Kymlicka, *Multicultural Citizenship*, p.160.

criteria for self-determined belief—can form a basis for activity that has value.”⁴⁹

We can now conclude that if Kymlicka’s interpretation of justice as fairness is right, then Rawls is a perfectionist instead of a neutralist.⁵⁰ However, this second-order conception has great difficulty in justifying congruence of the right and the good. This difficulty involves two incompatible conceptions of freedom stemming from Kant’s moral philosophy. The following session will focus on this issue.

6. Neutral Freedom and Good Freedom

The notion of self-determination stipulates that we should be free to deliberate and to choose our ways of life. It manifests a conception of person as free agency. According to Frankfurt, the very concept of a person consists in having desires of the second order about the first-order desires. Unlike animals, persons can form the second-order volitions and make choice according to their will.⁵¹ A free agent has the capacity for reflective self-evaluation of his desires and beliefs. In view of this, freedom of choice enables us to manifest our identity as free agent. Rawls endorses this conception of agency when he stresses that a free person is able to form, examine, and revise his conceptions of the good. Since persons have capacity to make free and rational choice, they are presumed to be responsible for their ends.⁵²

⁴⁹ Barry, *Justice as Impartiality*, p.132.

⁵⁰ Rawls expressly approves Kymlicka’s interpretation as “on the whole satisfactory” although “modulo adjustments that may need to be made to fit it within political liberalism as opposed to liberalism as a contemporary doctrine.”(PL:27) This indirectly confirms my claim that both early Rawls and Kymlicka have adopted a comprehensive liberal conception of the good.

⁵¹ Harry Frankfurt, *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), pp.11-25.

⁵² Rawls, *Collected Papers*, pp.369-70.

This conception of agency seems to imply that our freedom is equally manifested in choosing between good and evil, as much as in choosing different conceptions of the good. The upshot is we make our choice in accordance with our second-order volitions. Self-determination itself does not prescribe what we should choose. We cannot conclude that a good voluntary action expresses a greater degree of freedom than an evil one. Sidgwick calls this “*Neutral Freedom*”—“freedom exhibited in choosing wrong as much as in choosing right.”⁵³

This conception of freedom poses a serious challenge to Rawls’s congruence project: If congruence depends upon a person’s higher-order interest in expressing his nature as free being, and if the exercise of freedom is neutral between right and wrong, how can the desire to express freedom necessarily move the rational person to honour the regulative priority of justice? Does he not equally realize his nature when he chooses to act unjustly after informed deliberation? As Sidgwick famously puts, “the scoundrel must exhibit and express his characteristic self-hood in his transcendental choice of a bad life, as much as the saint does in his transcendental choice of a good one.”⁵⁴ Following the neutral conception, the freedom of choice and compliance with justice as fairness do not appear to necessitate each other. Nor can Rawls say that people who act unjustly are unfree. In short, the commitment to neutral freedom (or self-determination) does not result in congruence.

One possible way to resolve this difficulty is to appeal to a more “positive” conception of freedom. This conception must be able to provide resources for Rawls to say that abiding by principles of justice is the *only* way to realize our nature as free person.

⁵³ Henry Sidgwick, *The Methods of Ethics* (Indianapolis: Hackett Publishing Co., 1981), seventh edition, p.513.

⁵⁴ Sidgwick, *The Methods of Ethics*, p.516.

A person is free in proportion as he realizes his nature. Therefore, although people have neutral freedom to make their choice, they realize their *true freedom* by acting on moral principles which express that designated end. This conception is what Sidgwick calls “*Good*” or “*Rational Freedom*”, according to which freedom consists in one’s obedience to rationality, or moral laws based on pure practical reason.⁵⁵ A person is heteronomous if he is moved to act by his non-rational desires.

When Sidgwick draws this distinction between neutral and good freedom, he is talking about Kant’s moral philosophy. He suggests that both conceptions can be found in Kant. When Kant has to connect the notion of freedom with that of moral responsibility and free will, he refers to neutral freedom. When he intends to prove the possibility of unconditional obedience to moral law as such without the intervention of sensible impulses, and to exhibit the independence of reason in influencing choices, he refers to Good Freedom.⁵⁶ Sidgwick points out that these two conceptions are incompatible with each other. One cannot be described as free in making choice while being unfree in making wrong decision. The conditions of exercising neutral freedom and good freedom are entirely different. To avoid this paradox, Sidgwick argues that Kant must drop either of them to make his use of freedom consistent.

Rawls is well aware of Sidgwick’s critique of Kant. (TJ:254-56/224-225 rev.) Surprisingly, I find him following Kant in employing both conceptions of freedom in his theory. We have already seen that in justifying the principle of liberty and its priority, and responsibility for our choice, Rawls’s conception of free person is defined by neutral

⁵⁵ Sidgwick, *The Methods of Ethics*, p.512.

⁵⁶ Sidgwick, *The Methods of Ethics*, p.513.

freedom. We are free when we choose our ends. But in his response to Sidgwick's criticism against Kant, Rawls turns to the conception of good freedom. He says:

Kant's reply [to Sidgwick] must be that though acting on any consistent set of principles could be the outcome of a decision on the part of the noumenal self, not all such action by the phenomenal self expresses this decision as that of a free and equal rational being. Thus if a person realizes his true self by expressing it in his action, and he desires above all else to realize this self, then he will choose to act from principles that manifest his nature as a free and equal rational being. (TJ:255/224 rev.)

Rawls does not deny that both conceptions of freedom appear in Kant's theory. However, he stresses that although we can freely choose and act on any consistent principles, these principles do not equally express our true self as free and equal beings. Only when we act from those that manifest our nature can we become truly free. Since our nature is independently defined, this conception of freedom must refer to good freedom. For Rawls, his principles of justice are the most adequate option consistent with our nature because it is the result of agreement in the original position:

The parties qua noumenal selves have *complete freedom* to choose whatever principles they wish; but they also have a desire to express their nature as rational and equal members of the intelligible realm with precisely this liberty to choose, that is, as being who can look at the world in this way and express this perspective in their life as members of society...Thus men exhibit their *freedom*, their independence from the contingencies of nature and society, by acting in ways they would acknowledge in the original position. (TJ:255-56/225 rev., emphasis added)

Both conceptions of freedom are mentioned in the above paragraph. Rawls's main argument can be summarized as follows.

1. We are by nature free and equal rational beings.

2. We have a regulative desire to realize our nature.
3. The parties of the original position have complete *freedom* to choose whatever principles of justice they wish. (Neutral Freedom)
4. They would choose Rawls's principles because they could most fully express their nature as *free* beings. (Good Freedom)
5. Therefore, acting unjustly is acting in a manner that fails to exhibit our true self as free being in the well-ordered society. In this sense, an unjust person is an unfree person.

To make this argument valid, Rawls must demonstrate that the parties in the original position must have a common desire to choose Rawls's principles to express their *good* freedom even though they have complete *neutral* freedom to make choice. The crucial question is whether both conceptions of freedom can coherently exist in this argument. First of all, if we adopt the conception of neutral freedom, there seems no necessity between freedom of choice and the derivation of Rawls's principles. The very idea of free choice implies that people can make different choices for different reasons. Even in the highly restricted original position, as some critics have pointed out, the parties may have reason to choose the principle of average utility.⁵⁷ Of course Rawls would argue that the acceptance of his principles "is the *only* choice consistent with the full description of the original position. The argument aims eventually to be strictly *deductive*." (TJ:121/104 rev., my emphasis) But if so, the idea of contract becomes redundant and the claim that people have freedom to choose is a fraud. By the same token, there is no ground for Rawls to claim that people are unfree if they decide to act on non-Rawlsian principles after the veil

⁵⁷ For example, John Harsanyi, "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory," in *The Philosophy of Rawls*, ed. by Henry Richardson (New York & London: Garland Publishing, Inc., 1999), pp.234-46. Also see my discussion in the previous chapter.

is lifted. Therefore, (4) and (5) cannot be inferred from (3).

If we look at Rawls's argument more closely, we will notice that the conception of neutral freedom plays no substantive role in the argument. Even if we drop (3), (4) and (5) can still be derived from (1) and (2). The whole argument is actually based on the conception of good freedom. Rawls's main idea can be put as follows: Since we have a higher-order desire to express our nature as free being, we will only act on a conception of justice that can manifest our good freedom most fully. Rawls's principles are thus deductively derived. If there is any difficulty in reaching this conclusion, Rawls could simply modify the description so that the parties would "choose" his principles unanimously.

We now see that two conceptions of freedom cannot be used consistently in the same argument. Rawls faces the same critique of Sidgwick against Kant. In particular, in his argument for congruence, Rawls must adopt the conception of good freedom. For if people manifest their freedom through acting on any principles they choose, there is no way to justify the good of the sense of justice. What Rawls needs to establish is that the exercise of the sense of justice must embody a shared final end, namely the realization of our nature as free being, which is a higher-order good that every rational agent has reason to pursue. But if the realization of human nature becomes a common end, does it not imply that justice as fairness is a teleological theory? This claim sounds implausible because it is well known that the aim of *Theory of Justice* is to develop a deontological conception of justice as fairness to replace utilitarianism which is the most prominent representative of teleological theory. However, I believe that this claim not only makes sense, but also properly accounts for Rawls's later philosophical turn to political liberalism. In the following section, I will attest to this claim.

7. Justice as Fairness as a Teleological Theory

A teleological theory consists in two components: “the good is defined independently from the right, and then the right is defined as that which maximizes the good.” (TJ:24/21-22 rev.) It is important to notice that here the good is referred to non-moral, independently identifiable value. This is because, according to Frankena, “the moral quality or value of something to depend on the moral value of whatever it promotes would be circular.”⁵⁸ Teleological theories then make the right dependent on the non-morally good. In order to know whether something (an action, a policy, or an institution) is just, one must first know what is good in the non-moral sense, and whether the thing in question can maximize, or intended to maximize the good. These non-moral goods can be variously identified with happiness, pleasure, human excellence, knowledge etc. We can then have different teleological theories such as ethical egoism, perfectionism, hedonism and utilitarianism. What they commonly share is the idea of maximization of the good. Rawls further thinks that this is the deepest appeal of teleological theories because it seems to embody the idea of rationality. “It is natural to think that rationality is maximizing something and that in morals it must be maximizing the good.” (TJ:24-25/22 rev.) By definition, deontological theory is the very opposite of teleology that it neither specifies the good independently of the right, nor interpret the right as maximizing the good. (TJ:30/26-27 rev.) Rawls believes that justice as fairness meets the requirement of deontology in these two respects.

⁵⁸ Williams Frankena, *Ethics* (Englewood Cliffs, New Jersey: Prentice Hall, 1973), second edition, p.14. It is J. H. Muirhead who first uses the teleological/deontological division to describe all ethical theories; see his *Rule and End in Morals* (Oxford: Oxford University Press, 1932). Rawls does not use the term “non-moral good,” but he stresses that the goodness of things is a separate class which is judged without referring to what is right. (TJ:25/22 rev.)

Rawls agrees with utilitarianism about the definition of the good as the satisfaction of rational desire, or the satisfaction of rational plans of life. (TJ:408/358-359 rev.) What he objects to is the utilitarian claim that the satisfaction of any desire has some value in itself which must be taken into account in deciding what is right. This is unreasonable because not all desires are legitimate. Racial discrimination in a mainly white society, for instance, may bring a lot of pleasure to the white. However, this kind of pleasure should not be counted with equal weight like other preferences in the utility calculations because it violates minorities' rights. No matter how much satisfaction derived from those preferences, they have no moral weight whatsoever. A reasonable political morality, Rawls argues, should be the concept of right prior to that of the good. It follows that the principles of justice put limits on which satisfactions have value. They should impose restrictions on what are reasonable conceptions of one's good from the outset. Hence, "in justice as fairness one does not take men's propensities and inclinations as given, whatever they are, and then seek the best way to fulfil them." (TJ:31/27 rev.) This is the first reason why justice as fairness is a non-teleological theory. The second reason is that the idea of maximization does not play any role in justice as fairness. Although the parties in the original position want to secure as many social primary goods as possible, they would choose a principle of equal liberty and restrict social and economic inequalities to the greatest benefits of the least advantaged. So, "there is no reason to think that just institutions will maximize the good." (TJ:30/27 rev.) Since both essential features of teleological theory do not apply to justice as fairness, it is therefore identified as a deontological theory.

Rawls's characterization of teleology and deontology has been widely accepted. His criticism of utilitarianism implies that all kinds of teleological theories are doomed to

failure. This is mistaken because Rawls's Kantian argument for congruence is a type of teleological theory which need not face the difficulties of utilitarianism. I want to make two points to support this claim. First, I will show that Rawls's argument is grounded on a teleological account of human nature which makes the concept of the good prior to that of the right. Second, I will argue that the idea of maximization is not a defining feature of teleological theories. In other words, a conception of justice grounded on a certain ideal of human nature without adopting the maximizing principle is still a teleological theory. Justice as fairness is such an example.

To validate my first point, Rawls's argument is stated as follows:

1. The good is defined by the satisfaction of rational desire.
2. It is rational for us to have a higher-order desire to express our nature as free and equal rational beings. (TJ:574/503 rev.)
3. To express our 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:253/222 rev.)
4. Rawls's two principles of justice would be chosen by the parties in the original position which characterizes our nature as free and equal rational beings.
5. Therefore, acting in accordance with the principles of justice is something that we have regulative desire to do. The desire to act justly and the desire to express our nature turn out to be the same desire. So, acting from our sense of justice is a regulative good for rational persons. (TJ:572/501 rev.)

This argument noticeably depicts a teleological outlook. What is right is defined by the full expression of our nature. Our nature is our *telos*. It can only be fully realized and perfected through the effective exercise of the sense of justice. This should not surprise us because we have already seen that Rawls is a perfectionist according to the Kantian

interpretation, and perfectionism is a type of teleology.⁵⁹ There are some distinct features of Rawls's teleology which allow Rawls to avoid those difficulties faced by utilitarianism.

First, Rawls's theory can avoid the problem of illegitimate preferences. It does not hold that the satisfaction of desires stems from the same source, and therefore ranks equally without any independent discrimination. Rather, it takes the desire to realize our nature as qualitatively different desire. As Rawls puts, "in order to realize our nature we have no alternative but to plan to preserve our sense of justice as governing our other aims." (TJ:574/503 rev.) It is a higher-order regulative desire. When other desires are in conflict with the sense of justice, the latter has absolute priority. Since the sense of justice is characterized by Rawls's principles of justice, so any preferences that violate the principle of equal liberty will be disallowed at the very beginning.

Moreover, the calculation of the greatest sum of satisfaction does not have any place in deciding what is right or wrong in Rawls's case. The principles of justice are justified by its realization of freedom and equality. The possibility that the loss of freedom for some is justified by a greater good shared by others will not arise because a right to equal liberties is grounded on the Kantian conception of person. This conception takes the separateness of individuals seriously. However, this is not a deontological view as Sandel describes: "it describes a form of justification in which first principles are derived in a way that does not presuppose any final human purposes or ends, nor any determinate conception of the human good."⁶⁰ For according to the Kantian interpretation, justice as

⁵⁹ Rawls holds that there are two variants of perfectionism. The first one is that it is the sole principle of a teleological theory; the second one is that it is one of several principles in an intuitionist theory. The principle of perfection is balanced against others by intuition. (TJ:325/285-286 rev.)

⁶⁰ Sandel, *Liberalism and the Limits of Justice*, p.3.

fairness is justified by appealing to a particular vision of human *telos*. Rawls repeatedly reminds us that his whole theory is grounded on a conviction that we have a higher-order interest to realize our free nature. His principles of justice are constructed to represent and realize this fundamental interest.⁶¹ The primacy of justice is founded on a final human end. Once the overriding importance of realizing our nature is granted, a teleological theory warrants the priority of justice.⁶²

Sandel argues that this kind of teleological liberalism will put the primacy of justice at risk. For “where the right is instrumental to the advancement of some end held to be prior, the denial of liberty for some may be justified in the name of an overriding good for others.”⁶³ This concern brings out the second contrast between Rawls’s teleology and utilitarianism. As Frankena describes, utilitarianism regards the right simply as a means to promote an independent and non-moral good. Virtuous or right actions are merely instrumental to the maximization of the good. When the claim of right conflicts with the aggregation of non-moral good, the former ought to give way to the latter. Rawls rejects this view by stating that acting justly is constitutive to the perfection of human nature. The desire to be a just person and the desire to express one’s nature move a person to act in the same way. Doing what justice requires is not something separated from our good. On the contrary, it is an intrinsic value essential to our well-being. So, the problem of sacrificing the right for the good will not happen in Rawls’s case because doing so is irrational.

⁶¹ For instance, in the revised edition of *Theory of Justice*, Rawls remarks that “the basic rights and liberties and their priority are there said to guarantee equally for all citizens the social conditions essential for the adequate development and the full and informed exercise of their two moral powers—their capacity for a sense of justice and their capacity for a conception of the good—in what I call the two fundamental cases.” (TJ:xii, rev.) This idea has been fully elaborated in his “Kantian Constructivism in Moral Theory,” *Collected Papers*, pp.303-58.

⁶² Sandel actually admits that it is possible to establish the moral priority of justice without recourse to deontology. Mill’s liberalism is a case in point. Sandel, *Liberalism and the Limits of Justice*, p.3.

⁶³ Sandel, *Liberalism and the Limits of Justice*, p.18.

Someone may ask if there is such a great contrast between Rawls's proposition and utilitarianism, is it proper to label both of them as teleological theories? Korsgaard suggests that they are teleological in a different sense. Rawls's position is closer to classical teleological ethics represented by Aristotle while utilitarianism is a type of modern teleological theory. Although both share the same view that the good is realized through virtuous action, "classical teleologists argue that virtue is identical with the best state of a human being, while modern ones argue that virtue promotes an independent, nonethical good."⁶⁴ Rawls belongs to the classical camp because first, acting justly is constitutive to the perfection of human nature, and second, the realization of our nature is experienced as a moral good. Korsgaard is unsatisfied with Frankena and Rawls's differentiation of ethical theories into teleological and deontological, and assigning classical Greek theories along with utilitarianism to the first category and Kant's theory the second. For this widely accepted distinction has not only obscured the fundamental difference between classical and modern teleological theories, but also neglected the similarity between Kant and classical theory.⁶⁵ Following Korsgaard's categorization, there is no doubt that Rawls is a classical teleologist.

It may also be argued that even if justice as fairness gives priority to the concept of the good, it is still deontological because it denies the idea of maximization. Nevertheless, it is wrong to view maximization as a defining feature of teleology. Extending the principle of rational choice for one man to society as a whole, and pursuing the greatest

⁶⁴ Christine Korsgaard, (1998). "Teleological ethics," in E. Craig ed., *Routledge Encyclopedia of Philosophy*. (London: Routledge, 1998), retrieved November 20, 2005, from <http://www.rep.routledge.com/article/L103>. Although Korsgaard does not mention Rawls in this essay, I believe that her account of classical view provides a reasonable interpretation of Rawls's congruence argument.

⁶⁵ A major collection of essays offering a serious challenge to the traditional distinction about ancient and modern ethics can be found in S. Engstrom and J. Whiting (eds.) *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, (New York: Cambridge University Press, 1996).

net balance of satisfaction summed over all the individuals is a feature of utilitarianism, but not of teleology. Other teleological theories need not adopt this requirement. Take Aristotle as an example. For Aristotle, happiness (*eudaimonia*) is what rational agents ultimately desire. To lead a happy life depends on the realization of our distinctive human function as rational being through the practice of virtues. But the idea of maximization does not play any role in his thought. Although a political community must provide a congenial environment to enable citizens to lead a good life, it has no duty to maximize different kinds of virtues among citizens. Nor does Aristotle hold that we should compromise someone's good for the sake of greater happiness of others. There is no such maximizing formula in judging when an action is virtuous or not.

There are at least three reasons to explain why a teleologist need not accept the concept of maximization. The first is concerned with the source of value. Maximization presupposes that all values can be measured in accordance with a common scale. Their qualitative difference is thus entirely neglected because they are supposed to be reducible to the ultimate good. This is value monism that utilitarianism presumes. However, this meta-ethical view about value is questionable. For a teleological pluralist who believes that values are ultimately incommensurable, the very idea of maximization does not make any sense in the first place. He could argue that while there is a diversity of intrinsically valuable goods, they can only be realized in their distinct spheres according to their internal logic.⁶⁶ When there is conflict between different types of reason, no conclusive reason is available that one reason is necessarily outweighed by another because of the

⁶⁶ For different accounts of value pluralism, see Isaiah Berlin, *Four Essays on Liberty* (Oxford: Oxford University Press, 1969), p.171; Thomas Nagel, *Mortal Questions*, (Cambridge: Cambridge University Press, 1979), pp.128-41; Michael Stocker, *Plural and Conflicting Values* (Oxford: Clarendon Press, 1990).

lack of a common commensurable criterion.

The second reason is related to the nature of value. According to Rawls's understanding of teleology, once a good has been independently defined, what we should do is to promote it as far as possible. Rationality dictates that we ought to prefer the better to the lesser, and ought to do what is best. What is best is to maximize the good. But there are some goods to which the idea of maximization may not appropriately apply. Friendship, for instance, is unquestionably an important good for our life. Should we have a duty to maximize friendship then? Provided that we should, this statement can be understood in two different ways. It could mean either that we should make every effort to make as many new friends as possible, or that we should perfect our relationship with existing friends. These two ways to promote friendship are not conceptually incompatible with each. We can do both at the same time. But they connote very different meaning of "more is better." The first way only takes numbers into account. If we care about the *quality* of friendship, what we should do is perhaps to devote more time and love to our friends, and pay more attention to their needs. From utilitarian point of view, there is nothing wrong with betraying his friends for a much larger number of new friends if this can boost the net balance of happiness. Friendship, however, cannot be exchanged this way. It is absurd to demand a person to give up a special relation with his friend at any time when the sums come in from the utility network. What is at stake is not the accuracy of calculation, but doing so will distort the very nature of friendship. True friendship involves loyalty, trust and commitment which cannot be overridden by utilitarian calculation.⁶⁷ In this case, a teleologist needs not accept maximization as a necessary

⁶⁷ This is similar to Williams's critique of utilitarianism as an attack on a person's integrity. "A Critique of

means to promote the goodness of friendship.⁶⁸

Finally, Rawls's account of human nature has a built-in element to counter the trend of maximization. We know that Rawls's major complaint about utilitarianism is that it does not take seriously the plurality and separateness of individuals by conflating the desires of all persons into a coherent system. As a consequence, the loss of freedom for some is morally justified by a greater good shared by others. However, since Rawls views freedom and equality as an essential property of our nature, we have a fundamental interest in protecting our equal liberties in deciding the principles of social cooperation for reciprocity. Maximization is unacceptable because it will impair our identity as free and equal being. In Rawls's case, therefore, it is exactly because we have a regulative desire to realize our nature that we will not adopt the concept of maximization.

To conclude, my above argument shows that maximization of the good is not a defining feature of teleology. What is essential to a teleological political theory is that it justifies principles of justice with reference to some final purpose or the realization of human nature.⁶⁹ This is the main reason I ascribe Rawls's argument for congruence to a view of classical teleology. If my argument is valid, the widely accepted dichotomy of teleology and deontology then collapses. Here we come to the last concern of this chapter, to wit, can this teleological theory offer a successful argument for congruence consistent with the desire-based deliberative rationality?

8 The Difficulty of Congruence

Utilitarianism," pp.116-7.

⁶⁸ I am indebted to Tang Siu-fu and Tang Wai-sang for this point.

⁶⁹ Korsgaard, "Teleological Ethics."

In this section, I will show that there is a deep tension between teleology and the desired based rationality. This tension results in the failure of congruence and urges Rawls to recast his theory as a political conception of justice.

The primary aim of congruence is to vindicate the good of the sense of justice from the point of view of deliberative rationality so that the priority of moral motivation can be secured. A conception of justice is sufficiently stable if “from the standpoint of individual, the desire to affirm the public conception of justice as regulative of one’s plan of life accords with the principles of rational choice.” (TJ:577/505 rev.) More importantly, when rational individuals judge the desirability of the sense of justice, they “assess their situation independently from the constraints of justice.”(TJ:399/350 rev.) Putting congruence in the context of the Kantian interpretation, what Rawls strives to establish is that it is always rational for members in a well-ordered society, who have full knowledge of their conceptions of the good, to regard acting justly as a supremely regulative good because doing so can effectively express their nature as free and equal rational being. Stability is grounded on a combination of rationality and teleology.

The crucial question is whether deliberative rationality can have such normative force to lead individuals to accept and act upon Rawls’s account of human nature. We know that a plan of life is rational when it satisfies what we most want after informed deliberation. Our good is determined by the plan of life that we would adopt with full deliberative rationality. Thus an individual would be acting rationally if he would be doing what, all things considered, he wants to do most. This is a desire-based means-end conception of rationality. Whether an action is rational depends on whether it is the best means to satisfy our given fundamental desires. However, rationality itself cannot dictate which end we should have in the maximal class of ends. The end is the result of our choice. Given that

persons are left free to make choice, it is inevitable that “individuals find their good in different ways, and many things may be good for one person that would not be good for another.” (TJ:448/393 rev.) It is a permanent feature of modern society that rationality is coexistent with a plurality of conceptions of the good.

This account of rationality poses a serious challenge to Rawls: if people do not care about justice after careful and informed deliberation, there is no ground for him to further argue that it is rational for them to do so. The internal logic of desire-based rationality refrains Rawls from claiming the universal acceptance of the sense of justice as a higher-order end that rational agents would share. Rawls is on the horns of a dilemma. On the one hand, rationality cannot decide what desires we ought to have. People’s ends are inevitably diverse in a free society. On the other hand, to support his argument for stability, Rawls needs to show that people will commonly endorse the sense of justice as a regulative good embedded in their plans of life. A deep tension exists between pluralism and the priority of the sense of justice. In order to justify congruence, Rawls has taken the second horn by arguing that regardless of their different aims and desires, rational people have a shared end to realize their nature as free and equal rational being. This desire has absolute priority over other desire. As Rawls explains,

Therefore, in order to realize our nature we have no alternative but to plan to preserve our sense of justice as governing our other aims. This sentiment cannot be fulfilled if it is compromised and balanced against other ends as but one desire among the rest. (TJ:574/ 503 rev.)

When Rawls appeals to this teleological account of human nature, he of course believes that he is making an objectively true statement. Thus to act upon his principles is equivalent to realizing our most fundamental good. It is rational for people to adopt a

Kantian conception of the good to lead their life. Rawls later patently acknowledges this assumption:

This is the premise that in the well-ordered society of justice as fairness, citizens hold the same comprehensive doctrine, and this includes aspects of Kant's comprehensive liberalism, to which the principle of justice as fairness belong. (PL:xlii)

Having acknowledged this assumption, however, it is inconsistent with deliberative rationality and the fact of pluralism. After all, "it is rational for members of a well-ordered society to want their plans to be different." (TJ:448/393 rev.) There is no way to expect that rational people would commonly accept the same Kantian worldview in a well-ordered society. Rationality parts company with teleology. The real problem is not that Rawls's account of human nature is unconvincing, or that Rawls overlooks the fact of pluralism. It is rather that the desired-based rationality is unable to support teleology in a pluralistic society. This internal inconsistency has left little room for congruence. Surprisingly, after going to great lengths, Rawls concedes the bounds of rationality by saying that:

To justify a conception of justice we do not have to contend that everyone, whatever his capacities and desires, has a sufficient reason (as defined by the thin theory) to preserve his sense of justice. For our good depends upon the sort of persons we are, the kinds of wants and aspirations we have and are capable of. It can even happen that there are many who do not find a sense of justice for their good. (TJ:576/504-05 rev.)

This is equivalent to admitting that for those who do not accept the Kantian conception of person, we cannot recommend the sense of justice as a good to them since they do not have sufficient reason to do what justice requires. They may still comply with justice for prudential reasons. But they will not deem the sense of justice a regulative good

of their plans of life. In other words, except endorsing the Kantian interpretation of human nature, there is no other reason to justify the goodness of justice. The possibility of congruence is subject to a rather bold assumption that the majority of people would accept a Kantian conception of the good. In a liberal society where people have different kinds of wants and plans, Rawls's hope for congruence is fundamentally unrealistic.

What then might Rawls say to those who find it not a good for them to act justly? Of course Rawls could not dismiss them as irrational. There is no room for Rawls to make such a claim. Against all expectations, Rawls's answer is that

It is, of course, true that in their case just arrangements do not fully answer to their nature, and therefore, other things equal, they will be less happy than they would be if they could affirm their sense of justice. But here one can only say that their nature is their misfortune...Under such conditions penal devices will play a much larger role in the social system. (TJ:576/504 rev.)

This charge against those rational non-Kantian people is inconsistent with Rawls's conviction of justice. If their failure in recognizing the good of the sense of justice grows out of their unfortunate nature, why should they be responsible for that? According to Rawls, one's natural endowment is the outcome of the natural lottery, and is arbitrary from a moral point of view. We know that this is the main moral ground for his difference principle. Following this logic, those suffering from unhappiness owing to a nature not of their choosing should not be penalized. On the contrary, they are even entitled to some kind of compensation.

Furthermore, as I have thoroughly demonstrated in the first chapter of this thesis, Rawls's retreat to penal device to ensure strict compliance has confused the distinction between moral stability and social stability. What Rawls aims to achieve is the

motivational priority of the sense of justice. If this project fails and a certain portion of members lack sufficient motive to act in accordance with justice as fairness, the employment of coercive force is necessary to secure social stability according to Rawls. However, in this case the use of state power for social stability actually indicates the failure of moral stability. The pursuit of congruence is not a means to social stability, but an independent consideration for the desirability of justice as fairness. When Rawls concedes that there are no sufficient reasons to convince every rational individual to see the good of the sense of justice, it is not a practical matter, but a matter concerning the justifiability of justice as fairness. Later Rawls is fully aware of this distinction:

To clarify the idea of stability, let us distinguish two ways in which a political conception may be concerned with it. In one way we view stability as a purely practical matter...As long as the means of persuasion or enforcement can be found, the conception is viewed as stable. But, as a liberal conception, justice as fairness is concerned with stability in a different way. Finding a stable conception is not simply a matter of avoiding futility. Rather, what counts is the kind of stability, the nature of the forces that secure it. (PL:142)

Only in this background can we understand why later Rawls stresses that the problem of stability is fundamental to political philosophy and why an inconsistency internal to Part III of *Theory of Justice* has urged him to make basic readjustments of his whole enterprise. “All differences are consequences of removing that inconsistency.” (PL:xviii) For example, Rawls has totally dropped the Kantian interpretation and the idea of congruence in *Political Liberalism*. While he remains the theory of deliberative rationality, he has introduced the idea of reasonableness to constrain the use of rationality. The idea of overlapping consensus is proposed to replace congruence to justify stability for democratic society characterized by reasonable pluralism. Therefore, to evaluate Rawls’s later philosophical development, the most important frame of reference is to consider whether

his political liberalism can provide a better justification for moral stability. This is what I am going to do in the next chapter.