

Chapter Three: Rawls's Conceptions of Objective and Subjective Freedom

I. Introduction

Chapter one claims that Rawls and Hegel both articulate conceptions of objective and subjective freedom. Furthermore, the chapter asserts that both authors present the immanent critique of the fully free social world, but that neither holds that the critique is ultimately fatal to the ideal. The second chapter of the dissertation begins to unpack these claims in the case of Hegel. Now this chapter turns to Rawls's political philosophy. It reconstructs John Rawls's specific conceptions of the general concepts of objective and subjective freedom and provisionally locates the immanent critique in his writings.¹

This chapter is organized in the following way. It begins with an introduction of Rawls's theory of justice, which is necessary background material for the discussion to follow. Then, the chapter reconstructs Rawls's conceptions of objective and subjective freedom for individuals and for societies. The most important sources for the reconstruction are Rawls's writings on the autonomy of individuals and on the "well-ordered society" (TJ 453). The central results for which this chapter argues are the following. First, individuals are objectively free if and only if they obey the laws of society that (1) has a just basic structure and (2) includes the justification of the principles of justice in its public culture. A society whose institutions have these characteristics is also said to be objectively free or to realize objective freedom. Second, individuals are subjectively free if and only if they know and accept the principles of justice, understand how the principles apply to the basic structure of society, and act from their acceptance. A society all of whose members meet this requirement is itself subjectively free, or realizes subjective freedom.

After a brief consideration of previous literature that addresses the Rawls-Hegel intersection, this chapter builds on Rawls's definitions to make several more arguments. First, it contrasts Rawls's and Hegel's conceptions of objective and subjective freedom. Second, the chapter sets the foundation for the dissertation's reconstruction of the immanent critique. This chapter first demonstrates that Rawls is concerned with the persistence in a society of the two forms of freedom. The demonstration is carried out by establishing that Rawls treats "stability" (TJ 454) as a problem, and then by analyzing what he means by that term. Next, the chapter summarizes the parts of Rawls's writings that make the causal arguments that support the immanent critique. In particular, the chapter sketches two different routes by which Rawls thinks the objective freedom of a society may threaten its subjective freedom. The dissertation's reconstruction of these problems draws on Rawls's writings on stability, but goes considerably beyond an analysis of the concept. The discussion in this chapter of the relevant texts paves the way for the fuller treatment of them carried out by the second half of the dissertation.

II. Basic components of Rawls's theory of justice

Utilitarianism, applied to the question of social justice, states that just social institutions are those that maximize the total welfare of the people they affect.² Utilitarianism's attraction lies in the simplicity and clarity of its main principle and the body of philosophical arguments that have been marshaled to justify that principle.³ In his first major work, the 1971 A Theory of Justice, Rawls aims to provide a serious alternative to the utilitarian answer to the question of social justice. Rawls defines the subject of social justice as the distribution of primary goods (§15) for the basic structure of society (§2). Rawls means by the basic structure of society "the major institutions" (7)

of the social world: “the political constitution and the principal economic and social arrangements“ (7). In turn, A Theory of Justice defines primary goods as “things that every rational man is presumed to want” (62), such as political and civil rights, and income and wealth.

The practical output of A Theory of Justice consists of two principles and two priority rules:⁴

First Principle

Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.

Second Principle

Social and economic inequalities are to be arranged so that they are both:
(a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and
(b) attached to offices and positions open to all under conditions of fair equality of opportunity (TJ 302).

The first priority rule explains that the first principle is lexically prior to the second (302). This rule means that basic liberties cannot be diminished in order to increase income and wealth. The second priority rule says that fair equality of opportunity, (b), is prior to the principle (a), termed the “difference principle” (303).⁵ Obviously this scheme is more complex than the single utilitarian standard. However, the statement of it is equally sharp, and its derivation, to which I now turn, is equally well developed.

In A Theory of Justice, Rawls famously shows how to derive the principles and the priority rules that govern their application. In developing the derivation, Rawls comments that his “aim is to present a conception of justice which generalizes and carries to a higher level of abstraction the familiar theory of the social contract found, say, in Locke, Rousseau, and Kant” (11). His own argument is that the two principles would be agreed on by parties in a choice situation that he calls the “original position” (118). The

original position is a contract-style situation that he later elaborates as “hypothetical” and “non-historical” (PL 2647n). By hypothetical, he means that it is a thought experiment designed to justify the principles of justice, not an actual historical event. By “non-historical,” he means that the situation, by virtue of the way it is characterized, could not actually take place in the human world.

What are the key features of the original position? The original position is characterized by the conditions of “the circumstances of justice, the formal constraints of the concept of right, the veil of ignorance, and the rationality of the contracting parties” (TJ 118). In separate sections of A Theory of Justice, Rawls explains the constraints on the choice of the parties: the veil of ignorance (§24) and the formal constraints of the concept of right (§23). And he explains how they choose, characterizing their rationality (§25). The most noteworthy feature is the veil of ignorance, which conceals from the parties knowledge of facts about themselves:

No one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like. Nor, again, does anyone know his conception of the good, the particulars of his rational plan of life, or even the special features of his psychology such as his aversion to risk or liability to optimism and pessimism. (137)

Intuitively speaking, the veil of ignorance prevents the parties from being biased while coming to an agreement. Limited by these constraints, the parties must agree on principles of justice from a short list (§21). The short list includes other principles of distribution, most notably the utilitarian principle.

Rawls’s arguments from the original position received much attention in the years that immediately followed the publication of A Theory of Justice. My dissertation will not examine these arguments. However, it is worth noting that, in 1981, Rawls slightly

modified the principles of justice themselves. Starting with the essay “The Basic Liberties and Their Priority,” the two principles now read:

1. Each person has an equal right to a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all.
2. Social and economic inequalities are to satisfy two conditions. First, they must be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they must be to the greatest benefit of the least advantaged members of society. (“Basic Liberties” 5)

Rawls makes these changes in response to criticisms by H.L.A. Hart.⁶ Rejecting the notion of “maximizing” the scheme of liberties (“Basic Liberties” 48-9), Rawls replaces the corresponding language in the first principle with the language of full adequacy.⁷ He maintains the priority rules as before.⁸ After 1981, the principles remain unchanged.⁹ My dissertation does not examine Hart’s argument and Rawls’s response to it. Since I consider the issue independent of the questions I do treat, I simply will refer to the 1981 formulation of the principles from now on.

III. Rawls’s Conceptions of Subjective and Objective Freedom Reconstructed

While the principles of justice are the practical core of Rawls’s political theory, the theoretical foundations of his work are two ideals, or “model conceptions” (CP 307).¹⁰ The model conceptions consist of the autonomous person and the “well-ordered society” (TJ 5). Explicating individual autonomy and the well-ordered society in this section will allow me to connect Rawls’s commitment to justice to his concern with freedom. The section proceeds in the following way. First, I present Rawls’s notion of individual autonomy. I use this concept to derive provisional definitions of the objective and subjective freedom of both individuals and a society. Next, I turn to Rawls’s own explicit social ideal, the well-ordered society. I argue that certain features of the well-

ordered society should be used to revise the provisional definitions of objective and subjective freedom. According to the revised definitions, a well-ordered society is objectively and subjectively free. The section ends with a summary statement of the final reconstructed definitions of the objective and subjective freedom of an individual within society and a society as a whole.

Throughout his career, Rawls presents autonomy as a characteristic of acting individual citizens. His central writings on the autonomy of citizens are contained in three short works from the early 1980s. These works consist of the 1980 Dewey Lectures, “Kantian Constructivism and Moral Theory”; the 1981 Tanner Lectures, “The Basic Liberties and their Priority”; and the 1982 essay, “Social Unity and Primary Goods.” On the one hand, according to these pieces, the discussion in A Theory of Justice is sometimes incomplete, sometimes incorrect, and generally presented in a fragmented fashion. On the other hand, I will claim that the later discussions in Political Liberalism do not significantly develop the definition of autonomy. This part of my exposition therefore focuses most of its attention on the middle works from the early 1980s, with shorter portions at the beginning and end on the earlier and later texts.

Rawls’s definitions of autonomy in A Theory of Justice occur in §40, “The Kantian Interpretation of Justice as Fairness,” and in §78, “Autonomy and Objectivity.” These sections reveal autonomy to be a variety of rational and self-determined action engaged in by individual citizens. The most succinct version of the definition appears in §78: “Acting autonomously is acting from principles that we would consent to as free and equal rational beings, and that we are to understand in this way” (TJ 516). Rawls holds that his two principles of justice are such principles.

The definition of autonomy has a subjective and objective component. Autonomous individuals do not merely act according to the right principles, they act *from* those principles, which in turn they *understand* correctly. In other words, autonomous action is self-determined and self-conscious. Rawls does not identify this aspect of the definition as subjective, but he does include it whenever he discusses autonomy: in §40, for instance, he says that autonomous actors act “knowingly” (253) or “consciously” (255). In §78, moreover, Rawls stresses: “[The] principles [of justice] are objective” (516). By that he means that they are the same for everyone, and justified to everyone by their derivation. Moreover Rawls thinks that the content of principles is designed to express the freedom of individuals when acted on because they are the “principles that [individuals] would acknowledge under conditions that best express their nature as free and equal rational beings” (TJ 515). The device of the original position ensures these conditions. Now, while this argument is discussed in §40 of TJ, it is amended and elaborated in “Kantian Constructivism.” Thus this paragraph will not dwell on it. The key point here is simply the connection to freedom that Rawls draws.

I move on to the more extensive discussions in the three essays from the early 1980s. In “Kantian Constructivism in Moral Theory,” Rawls explicitly connects autonomous action to a set of capacities for action. He elaborates the model-conception of the person, which is characterized by two capacities, or “moral powers” (CP 312) and two “corresponding highest-order interests” (312) in exercising these capacities. The first moral power is a capacity for a sense of justice (312). Exercising one’s capacity for a sense of justice includes understanding the principles of justice, applying them to institutions, and acting from reasons of justice (312). The second moral power is the

capacity to “form, revise, and rationally to pursue a conception of the good” (312). Conceptions of the good are encompassing worldviews that specify what ends are worthy of pursuit over a human lifetime. Conceptions of the good include religious doctrines such as Christianity, Judaism, or Hinduism. They also include philosophical doctrines of human flourishing, such as utilitarianism, aestheticism, or hedonism. Forming and revising such views are contemplative pursuits, while rationally pursuing them requires action in the world. Finally, Rawls assumes that the moral person has a single, determinate conception of the good at any one time, and a higher-order interest in pursuing it (313). By subordinating the moral person’s determinate conception of the good to the ability to revise that conception, Rawls makes it clear that the moral person’s identity is not tied to any particular conception of the good. An individual may convert from Christianity to Judaism to Buddhism over the course of a lifetime, and still remain the same person from Rawls’s moral perspective.

In “Kantian Constructivism,” Rawls distinguishes between two forms of autonomy. The moral person is both “rationally autonomous” (CP 308) and “fully autonomous” (320). Rational autonomy, Rawls says, “roughly parallels . . . the notion of rationality found in neoclassical economics” (308). However, the individual as rationally autonomous¹¹ is not concerned with just anything; he is concerned with the pursuit of his conception of the good. In “The Basic Liberties and Their Priority,” Rawls says: “rational autonomy is acting solely from our capacity to be rational and from the determinate conception of the good we have at any given time” (20). He is also concerned with his highest-order interests, insofar as these promote or constitute his determinate conception of the good (“BL” 29). Rawls argues that the rationally

autonomous person will take into account both highest-order interests (“BL” 24-39). It is worth emphasizing that Rawls places significant constraints on the ends of the rational autonomous individual. Some accounts of instrumental rationality place far fewer constraints on ends. Thus rational autonomy is a morally richer notion than such minimalist versions of instrumental rationality.

In both “Kantian Constructivism” and in “Social Unity and Primary Goods,” Rawls elaborates his description of the primary goods, connecting them conceptually to the model conception of the person and his capacity for rational autonomy. Recall from above that the person is moral and has highest and higher-order interests. Now, primary goods are what any such person would want given these highest and higher-order interests (though what conception of the good the person has is not specified) (CP 314). Rawls says that primary goods are “social background conditions and general all-purpose means normally necessary for developing and exercising the two moral powers and for effectively pursuing a conception of the good” (313). More concretely, they include “the basic liberties” (313), “freedom of movement and free choice of occupation” (313), “powers and prerogatives of office and positions of responsibility” (313), “income and wealth” (314), and “the social bases of self-respect” (314). For example, Rawls assumes that a Christian and a follower of Ayn Rand’s objectivist philosophy would both value more rather than less money to advance their respective ends. Access to primary goods enables the rational autonomy of citizens, whatever their determinate conceptions of the good may be.

Now I move to full autonomy. This definition requires mention of the principles of justice. In “Kantian Constructivism,” Rawls writes: “For it is by affirming the first

principles that would be adopted in this situation and by publicly recognizing the way in which they would be agreed to, as well as by acting from these principles as their sense of justice dictates, that citizens' full autonomy is achieved" (CP 315). This definition is the same as the definitions of autonomy offered in A Theory of Justice. Once again the subjective and objective aspects of autonomy are both present. In this piece, Rawls also refers to the principles, by virtue of their objective content, as "reasonable" (317).¹²

"Kantian Constructivism" presents a lengthier account of the connection of the content of the principles of justice to freedom; indeed, that is the piece's main theoretical goal. A complete explication of Rawls's "constructivism" (303-4) will wait until chapter five. Here my aim is to give some minimal background. Rawls sets the stage for his later constructivist theory in §40 of A Theory of Justice. In this discussion, Rawls states that "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). Note that, while "expression" of one's nature requires the subjective affirmation of certain principles, it also requires that the principles have the appropriate objective content. In §40, Rawls makes a second claim: "By acting from the [the] principles [of justice] persons express their nature as free and equal rational beings subject to the general conditions of life" (TJ 252-3). This claim relies on the previous definition of expression, combined with Rawls's contention that the principles, Rawls thinks, are what a free person would will. Whereas §40 only sketches the argument for this assertion, "Kantian Constructivism" refines and elaborates it. In "Kantian Constructivism," Rawls claims that constraints on information and the characterization of the parties are the means by which the principles are connected to the model conception of the person. The priority of

the first principle to the second reflects the priority of the highest-order interest of the moral person in having and revising a conception of the good to the higher-order interest of pursuing a determinate such conception.

Now I turn to the writings on autonomy collected in Political Liberalism. In these works, Rawls emphasizes that the concepts involved are “political” (PL xliv) rather than “comprehensive” (xliv). He now distinguishes his notion of autonomy from “moral autonomy expressed in a certain mode of life and reflection that critically examines our deepest ends and ideals, as in Mill’s ideal of individuality, or by following the best one can Kant’s doctrine of autonomy” (xlv). Rawls alters his terminology accordingly. The moral person is now referred to as the “citizen” (xlv). The term “autonomy” is prefaced, implicitly or explicitly, with the modifier “political” (xlv, 402). The theory of Kantian constructivism is reinterpreted as “political constructivism” (89) while the primary goods are characterized as “citizens’ needs” (187).

While Political Liberalism marks a major change in Rawls’s political theory, it is important to locate that change carefully. This dissertation will follow previous commentators in arguing that the change represented by Political Liberalism lies in its account of stability,¹³ a concept that I introduce below. In contrast, the labeling of autonomy as “political,” in the absence of any alteration in the definition of autonomy itself, is not a major alteration to the theory. Rawls’s discussions of autonomy in Political Liberalism more or less repeat the earlier discussions. In fact, one of the essays in Political Liberalism, “The Idea of Public Reason,” incidentally offers the clearest statement in his writings of the importance of the subjective dimension of full autonomy. Rawls writes:

The government's authority cannot . . . be freely accepted in the sense that the bonds of society and culture, of history and social place of origin, begin so early to shape our life and are normally so strong that the right of emigration (suitably qualified) does not suffice to make accepting its authority free, politically speaking. Nevertheless, we may over the course of life come freely to accept, as the outcome of reflective thought and reasoned judgment, the ideals, principles, and standards that specify our basic rights and liberties, and effectively guide and moderate the political power to which we are subject. This is the outer limit of our freedom. (222)

He elaborates in a footnote: "Here I accept the . . . view that what we affirm on the basis of free and informed reason and reflection is affirmed freely; and that insofar as our conduct expresses what we affirm freely, our conduct is free to the extent it can be" (222-3 9n). Here Rawls underscores that the reflective acceptance of rational principles is necessary for the complete freedom, or full autonomy, of the individual citizen. The emphasis is worth noting, since usually he does not underscore the subjective dimension of individual full autonomy.

In sum, the pieces of Rawls's final account of autonomy include the model-conception of the citizen; the distinction between rational and full autonomy; the conceptual connection of the citizen to both kinds of autonomy; and the placement of primary goods in this scheme. Drawing on these discussions of autonomy, I now can offer a provisional reconstruction of the Rawlsian conception of objective and subjective freedom.

I begin by using Rawls's definition of full autonomy to reconstruct definitions of the objective freedom of the individual citizen and then of society. According to the general concept of individual objective freedom, an individual citizen is objectively free only if he enables his individual freedom when he obeys the laws of his social institutions. I contend that a fully autonomous citizen within a just society is objectively

free. By Rawls's definition of full autonomy, he obeys the principles of justice, and therefore the laws of a just basic structure. Now, a just basic structure provides primary goods to the citizens of the society. We have seen that primary goods enable the rational autonomy of individuals. Moreover, according to the theory of constructivism, an individual who knowingly obeys the laws of a just basic structure expresses his identity as a free citizen. This expression depends on the fact that institutions produce a particular objective distribution of primary goods. So, such individuals causally *produce* the conditions of their rational autonomy according to a distribution that is necessary to *express* their identity as free. We may say that a citizen who obeys the laws of a just basic structure enables his individual freedom in these two ways. Therefore, I conclude that such a citizen is objectively free. Next, chapter one says that a society as a whole is objectively free only if its institutions, generally complied with, enable the individual freedom of all its citizens. By the same reasoning just carried out, a society with a just basic structure is objectively free.

The preceding paragraph argues that Rawlsian justice, once unpacked, is a conception of objective freedom. Next I reconstruct Rawls's notions of the subjective freedom of the individual and of a society. An individual is subjectively free, according to the general concept, if and only if he affirms the institutions of his social world as his own and obeys its laws and norms from that affirmation. I now contend that a fully autonomous individual in a just society is subjectively free. A fully autonomous citizen does not merely obey the principles of justice but acts *from* the principles of justice and understands and affirms them. Subjective understanding, affirmation, and corresponding action belong to the definition of full autonomy. In a just society the principles of justice

are embodied in the laws of the basic structure of society. So, Rawlsian citizens in a just society who knowingly affirm and act from the principles of justice thereby knowingly affirm and act from the institutions of their society. Therefore we may say that individuals are subjectively free if and only if they knowingly affirm the principles of justice and act from them. Next, recall that a society realizes subjective freedom if and only if all its members are subjectively free. We may also say that a just society all of whose citizens knowingly affirm the principles of justice and act from those principles is subjectively free.

With these comments, I conclude the provisional reconstruction of objective and subjective freedom. Now I turn to Rawls's notion of a "well-ordered society" (TJ 5). Like the model conception of the citizen, a well-ordered society is an ideal; in this case, an ideal of the social world. The discussions of this notion make a continuous thread that runs through his writings. In this section, my goal is to show that a well-ordered society is a social world that realizes objective and subjective freedom. To make that argument, I start with the provisional definitions above. However, I also use the discussion of a well-ordered society to motivate revisions in the definitions of both subjective freedom and objective freedom of individuals and society. These revisions are not arbitrary attempts to produce a "fit," but are based on the general concepts of objective and subjective freedom and their derivation from the interpretation of freedom as rational self-determination.

In §69 of A Theory of Justice, Rawls defines a well-ordered society as one "in which everyone accepts and knows that the others accept the same principles of justice, and the basic social institutions satisfy and are known to satisfy these principles" (TJ

454). Moreover, these facts have a clear connection to the citizens' actions: "A well-ordered society is also regulated by its public conception of justice. This fact implies that its members have a strong and normally effective desire to act as the principles of justice require" (454). This definition remains constant throughout his writings. For instance, in Lecture I of Political Liberalism, "Fundamental Ideas," Rawls says that a well-ordered society "is a society in which everyone accepts, and knows that everyone else accepts, the very same principles of justice; and second . . . its basic structure . . . is publicly known, or with good reason believed, to satisfy these principles" (35). Furthermore, he continues, "its citizens have a normally effective sense of justice and so they generally comply with society's basic institutions, which they regard as just" (35). Similar passages occur in "Kantian Constructivism" (CP 308-9) and Justice as Fairness (8-9).

Note that the well-ordered society has two explicated stated features. First, it is a just society. Second, all of its members regard it in a certain way and act in a certain way. All of them affirm the principles of justice, know that the basic structure meets the principles, and act from that knowledge and acceptance. Throughout the rest of my dissertation, I refer to the second aspect of the definition as the "unanimity condition." Now, this choice of terminology may seem misplaced, since Rawls generally speaks of unanimity as a constraint on the choice of principles in the original position (TJ 122). However, at one point in A Theory of Justice, after noting that "the original position is so characterized that unanimity is possible; the deliberations of any one person is typical of all," Rawls adds that "the same will hold for the considered judgments of the citizens of a well-ordered society effectively regulated by the principles of justice" (TJ 263). The

second clause bears out my use of the term “unanimity” as applied to a well-ordered society.

Now I consider whether the well-ordered society is objectively and subjectively free, according to my provisional definitions of these terms. First, I remark that the well-ordered society is indeed objectively free, since it is just. But is it subjectively free? Recall that a society is subjectively free if and only if all of its members are. So we must examine the subjectivity of the members of the well-ordered society. Rawls says that in a well-ordered society all citizens accept the principles of justice and act from that acceptance. In “Kantian Constructivism and Moral Theory,” Rawls says: “In a well-ordered society we stipulate that the justification of the principles of justice as the outcome of the original position is publicly understood” (CP 320). This full justification “includes connecting the moral doctrine’s model-conceptions with the society’s particular conception of the person and of social cooperation (325). However, a few pages later, Rawls restates and then explicitly modifies his position:

I suppose this full justification to be publicly known or, better, at least publicly available; this weaker condition allows for the possibility that some will not want to carry moral reflection so far, and certainly they are not required to do so. (325)

The availability of the full justification of the two principles is officially characterized as part of the “publicity” of a well-ordered society.¹⁴ In Political Liberalism, in the lecture entitled “The Powers of Citizens and their Representation,” Rawls repeats the point:

I suppose this full justification also to be publicly known, or better, at least to be publicly available. This weaker condition (that full justification be available) allows for the possibility that some will not want to carry philosophical reflection about political life so far, and certainly no one is required to. (67)

Finally, the 2001 Justice as Fairness: A Restatement, Rawls’s last presentation of the theory of justice before his death in 2002, includes essentially the same statement (121).

My provisional definition of individual subjective freedom is based on Rawls's notion of full autonomy. According to this definition, a subjectively free citizen's acceptance of the principles of justice is knowing or conscious—that is, fully rational. By this account, the members of a Rawlsian well-ordered society, who simply affirm the principles, fall short of subjective freedom. However, chapter two uses Neuhaus's writing to argue that a Hegelian social member need not be fully rational to count as subjectively free, as long as his action meets certain minimum standards of rationality. We can extend this thought to Rawls's political theory and revise the definition of subjective freedom accordingly. According to the revised definition, Rawlsian citizens are subjectively free if and only if they affirm the principles of justice and act from that affirmation. They need not master the full justification. A society is subjectively free if all of its citizens meet those requirements. According to this new definition, a well-ordered society realizes subjective freedom.

This alteration in my definition of subjective freedom is matched by an alteration in the definition of objective freedom. Recall that the full justification of the principles is present in the public culture of a well-ordered society, “reflected in law and public institutions, and in the philosophical and historical traditions of their interpretation” (CP 325). As I will explain later in this chapter, the subjective freedom of reflective individuals ultimately depends on their opportunity to learn the full justification of the principles. In “Kantian Constructivism,” Rawls explicitly says that the public availability of the justification of the principles “is a precondition of freedom” (326) for the citizens of a well-ordered society. Now, the concept of objective freedom says that a society realizes it only if that society's institutions enable the individual freedom of each citizen.

Because the public availability of the principles' justification enables the subjective freedom of individuals, I add it to the definition of Rawlsian objective freedom for societies and individuals. In other words, a society realizes objective freedom if and only if it is just and the full justification of the principles of justice is publicly available.

Individuals are objectively free when they obey the laws of such a society.

Finally, I will use the definition of the well-ordered society to add one last requirement to the definition of subjective freedom for individuals, and, by extension, for societies. Notice that Rawls says that all members of a well-ordered society know that the basic structure satisfies the principles of justice. Rawls never explicitly says why he includes this requirement in his definition. However, if an individual is to endorse his basic institutions on the grounds that they satisfy the principles of justice, he must of course know that they do satisfy the principles. It is a good assumption that Rawls includes this requirement because it is a condition for an individual's reasonable endorsement of his basic institutions. Even if that endorsement is not fully reasonable, if it meets the requirement in question, it is not merely capricious. Now, the concepts of subjective and objective freedom are based on the prior notion of freedom as rational self-determination. To know that the basic structure is just is a feature of the subjectivity of a member of a Rawlsian well-ordered society. Such knowledge is a condition on the rationality's of the agent's action. Therefore, it is an appropriate addition to the definition of Rawlsian subjective freedom. The final definition of Rawlsian subjective freedom is then the following. Individuals are subjectively free if and only if they accept the principles of justice, know that these apply to the basic structure of society, and act

from their acceptance and knowledge. A society is subjectively free if and only if all of its members meet these requirements.

My reconstruction of Rawls's conceptions of objective and subjective freedom is now complete. Like Hegel's conceptions of objective and subjective freedom, Rawls's reconstructed conceptions are grounded on his interpretation of freedom as self-determined action. Both accounts moreover assume that there is more than one kind of self-determined action, and that self-determined actions of whatever type are engaged in by entities with certain capacities and in certain environments. The institutions of the state are instrumental to providing for self-determined actions of the lower, less perfect type. In Rawls's case, the basic structure enables rational autonomy; in Hegel's, the institutions of *Sittlichkeit* enable personal and moral freedom. Obeying the laws and norms of these institutions constitutes a second, higher form of freedom: in Rawls's case, full autonomy, and in Hegel's case, what Neuhausser calls social freedom. Finally, a complete account of the institutions of social life includes a description of public institutions that work to ensure subjective freedom.

IV. Other Treatments

At this point, the dissertation has presented two conceptions of subjective and objective freedom, Hegel's and Rawls's and shown how each is built on the foundation of freedom as rational self-determination. These conceptions in turn can be examined side by side and contrasted in detail along two dimensions. However, it is worth pausing first to examine the existing literature on the Rawls-Hegel intersection and to consider whether or not these works correctly describe the shared features of the two authors' treatments of freedom.

Some pieces note aspects of Rawls' and Hegel's treatments, but do not offer a complete account of the topic. For instance, in his 1994 article, "A More Democratic Liberalism," Joshua Cohen recognizes that Rawls's notion of individual freedom has a subjective (1516-7)¹⁵ and objective (1517) dimension. Moreover, he shows that Rawls bears a relationship to Hegel, and explains that Hegel's philosophy "link[s] the expression of our free nature to the institutions of a state whose aim is the realization of the good—understood as the expression of our nature [1]" (1508). However, he does not mention that Hegel also views endorsement of the state by its members as crucial to the members and the state's realizing freedom. Stephen Houlgate's 2001 essay "Hegel, Rawls, and the Rational State," on the other hand, omits mention of the subjective dimension of freedom in the case of both authors. Houlgate emphasizes the connection between Hegel's rational state and Rawls's well-ordered society. He notes: "Like Hegel's state, Rawls's well-ordered society is held together, not by mere force or economic power, but by publicly recognized principles and laws" (250). Addressing the content of the laws, Houlgate says that "Hegel and Rawls . . . agree that the principles of justice governing the modern state should be founded on the concept of rational freedom" (250). However, Houlgate does not mention that recognition of the laws is itself an aspect of freedom for both Rawls and Hegel.

Houlgate and Cohen's treatments of Rawls and Hegel are not inaccurate, just incomplete as representations of the two authors' thoughts on freedom. In contrast, in her essay "Hegel's Critique of Rawls," J.B. Hoy does substantively misrepresent Rawls's conception of freedom by leaving out a key part of it. While Hoy displays a

comprehensive understanding of Hegel's writing on freedom, she equates Rawls's whole doctrine of freedom with his notion of rational autonomy. She writes:

On Hegel's view freedom of subjectivity entails the interdependence of persons and the fulfillment of individuals in the laws, institutions, and customs of their societies. Only by participation in the rational ethical substance, in *Sittlichkeit*, is individual freedom fully realized (*Rph*, #'s 142-155). Rawls's emphasis on equal basic liberties geared to protecting the individual from interference and restriction thus misses the point of Hegel's understanding of freedom as the principle of modern society. (418)

This understanding leaves out the fact that acting from the principles of justice is fully autonomous action, a second kind of freedom distinct from the rational autonomy enabled by the basic liberties.¹⁶

Two previous treatments do exist that cover both dimensions of freedom for both authors. In her 1991 paper, "Rawls, Hegel, and Communitarianism," Sibyl Schwarzenbach maps Rawls's two concepts of autonomy onto Hegel's notions of freedom. Of Hegel and freedom, Schwarzenbach writes:

The important point is that for a person to obtain "substantive freedom" in Hegel (and not just Willkuer or choice), it is not enough simply to overcome the alien and compelling character of the natural physical world (achieved through labor and property), nor is it enough to bring order to the chaotic inner world of one's desires (achieved primarily through moral reflection); one must also overcome the compulsory nature of human, communal life. The latter is achieved, in Hegel's view, through educating oneself to an awareness of universal ends as well as by participating in the construction of the rational character of public social life (PR, paras. 149, 260). "Substantive freedom" can only be fully instantiated, he claims, with a community wide "reciprocal recognition" of freedom as "lived social practice." The latter Hegel terms *Sittlichkeit* or "rational ethical life." (553)

This paragraph alludes to the three forms of freedom presented by Hegel in the Philosophy of Right: personal freedom, moral freedom, and social or substantive freedom. Schwarzenbach's description of substantive freedom refers to its objective and subjective dimensions, although she does not label them as such. Then, in contrast to

Hoy, Schwarzenbach does go on to note that Rawls's defines two distinct notions of autonomy, rational and full (554). Her statement suggests that full autonomy in Rawls's thought corresponds to social or "substantive" freedom in Hegel's theory (554).

Finally, Julius Sensat's 2003 piece "Classical German Philosophy and [G.A.] Cohen's Critique of Rawls" includes a description and comparison of Rawls and Hegel on freedom that is basically complete. Although he does not use the terms "subjective" and "objective," Sensat correctly states:

For Hegel, a philosophical account of existing social arrangements explains how they actualize the concept of freedom. As noted above, such an account can contribute to that very actualization, by itself taking on a social role in the public culture and thereby completing the requirements of the rational functioning of those arrangements. It completes the actualization of the concept of freedom by making the social world more self-consciously rational. (Sensat 321)

Here, Sensat correctly remarks that, for Hegel, subjective endorsement of the state by its members is necessary if it is to realize freedom. This comment refers to the state as a whole in Hegel's philosophy. Then, later, Sensat notes that Rawls shares Hegel's belief in the "need for a social world in which citizens autonomously maintain the social conditions of their autonomy" (340). The first "autonomously" in this sentence points out the subjective dimension of individual freedom. Furthermore, citizens who maintain the "social conditions of their autonomy" are by definition objectively free. While this account is not embedded in a discussion of Rawlsian full autonomy or Hegelian substantive freedom, the comparison as it stands is still worthy of remark. Sensat correctly points to both dimensions, subjective and objective, of fully free action, and shows that Rawls and Hegel value such action.

V. Contrasting the two conceptions

Sections III and IV of this chapter emphasize the common philosophical elements of Rawls and Hegel's political theories. Their interpretation of freedom as rational self-determination is the most fundamental common element. From that notion follows conceptions of the freedom of the individual as a member of society and of the social world as a whole. Each of these notions, moreover, has an objective and subjective dimension. This section will shift the emphasis of my treatment and contrast the two authors' conceptions of the objective and subjective freedom of the citizen and of the social world. This section argues that Rawls's conceptions are more individualistic and more egalitarian than Hegel's. The arguments here are fairly brief.

Chapter two of the dissertation follows Neuhauser in arguing that, for Hegel, the social world realizes objective social freedom in two ways. First of all, it realizes objective freedom by providing the social conditions for the individual freedom of its members. Second, Hegel claims that *Sittlichkeit* as a whole realizes a form of autonomy not attributable to its individual members and more complete than individual autonomy. The two forms of objective freedom are related, because realization of the first form of objective social freedom is necessary for the realization of the second. Still, the second is a significant addition that gives Hegel's political theory its distinct flavor.

Neuhauser calls the second conception of objective freedom strongly holistic, by which he means that it is attributable to the social whole but not to the individuals within it. In contrast, no aspect of Rawls's conception of objective freedom is strongly holistic. The reconstruction that this chapter performs is founded on the objective portion of Rawls's notion of individual autonomy. Here, individuals objectively promote their own freedom by complying with just institutions. Now, I do modify the definition of

objective freedom after the introduction of the concept of the well-ordered society.

However, at no point does this chapter's account introduce a feature of the social world that could be said to constitute its freedom and is not a condition of individual freedom.

Because Rawls's conception of objective freedom does not have a strongly holistic component and Hegel's does, it is more individualistic than Hegel's. This claim simply follows conceptually from the definition of a strongly holistic property. The introduction of the new terminology does not add to the conceptual resources that the dissertation draws on, but enlarges the available forms of expression.

This section now contrasts Rawls and Hegel's conceptions of subjective freedom, and their conceptions of objective freedom as those relate to subjective freedom. As chapter two demonstrates, Neuhausser argues that Hegel thinks that a citizen's subjective freedom consists in an attitude of affirmation toward his social institutions (Neuhausser 82-113). This attitude may be, but need not be, deeply reflective. Next, according to my reconstruction of Rawls's conception of subjective freedom, the citizens of a well-ordered society should (1) know and accept the principles of justice; (2) recognize that the basic structure realizes the principles; and (3) act accordingly. Again, citizens are not required to master the full justification of the principles of justice. Inasmuch as neither author requires subjectively free citizens to be fully rational in their affirmation of their social world, the two accounts appear identical.

Still, the dissertation will argue that there is a fundamental difference between Hegel's and Rawls's accounts of the attitudes of social members toward their institutions. Chapter two notes that Hegel claims that the social world should be, in principle, fully intelligible to all its members. However, as Neuhausser points out, Hegel has different

expectations of different, identifiable, segments of the citizenry. Hegel offers different routes by which he expects that they will achieve subjective social freedom. Chapter eight of the dissertation will argue that, by examining *Sittlichkeit* closely, it turns out that Hegelian *Sittlichkeit* does not actually offer all of its members access to the full account of its objective freedom. On the other hand, chapter eight argues that Rawls's well-ordered society does offer all of its citizens access to the full justification of the principles of justice, and the explanation of how the principles apply to the basic structure. Consider a measure of the rationality of a citizen's affirmation of his basic institutions. While Rawls does not guarantee equality of outcomes for this measure, he, unlike Hegel, does wish to ensure equality of opportunity. Because of his commitment to this form of equality of opportunity, Rawls's conceptions of objective and subjective freedom are more egalitarian than Hegel's.

This section has shown that Rawls's conceptions of objective and subjective freedom are more individualistic and more egalitarian than Hegel's.¹⁷ The dissertation will return to and enlarge these arguments in chapters six and eight, respectively. Right now, we can put the difference between Rawls and Hegel aside and turn to the critical material in Rawls's political theory.

VI. Stability, Unanimity, and Freedom

This section introduces Rawls's writings on stability and connects them to objective and subjective freedom. Rawls discusses stability in all of his major works. He uses the term in three distinct but related ways. He speaks of the stability of the principles of justice, the stability of a well-ordered society, and the stability of a society "for the right reasons" (PL xlii). Rawls does not take any form of stability for granted,

but instead regards each as potentially lacking. I will argue that, in doing so, Rawls calls into question the persistence of a fully free society.

In A Theory of Justice, Rawls introduces the theme of stability in §69, “The Concept of a Well-Ordered Society.” The parties in the original position must check that the conception of justice they have chosen is stable. In this section, Rawls presents stability as a property of the correct conception of justice. He writes:

Now a well-ordered society is also regulated by its public conception of justice. This fact implies that its members have a strong and normally effective desire to act as the principles of justice require. Since a well-ordered society endures over time, its conception of justice is presumably stable: that is, when institutions are just (as defined by this conception), those taking part in these arrangements acquire the corresponding sense of justice and desire to do their part in maintaining them. (454)

In his 1975 essay, “The Independence of Moral Theory,” Rawls again treats stability as a property of a conception of justice. A conception’s stability depends on “whether its principles generate their own support in a society, or social group, in which these principles are publicly realized” (CP 294).

At other points, Rawls speaks of stability as a property of a well-ordered society. For instance, in §69 of TJ, he notes that stability can be considered a property of systems. He then says:

The relevant systems here, of course, are the basic structures of the well-ordered societies corresponding to the different conceptions of justice. We are concerned with this complex of political, economic, and social institutions when it satisfies, and is publicly known by those engaged in it to satisfy, the appropriate principles of justice. (TJ 457)

In “Kantian Constructivism and Moral Theory,” he again treats stability as a property of a well-ordered society, explaining:

The stability of a well-ordered society is not founded merely on a perceived

balance of social forces the upshot of which all accept since none can do better for themselves. To the contrary, citizens affirm their existing institutions in part because they reasonably believe them to satisfy their public and effective conception of justice. (CP 324)

It should be clear from Rawls's definitions that the second sense of stability is logically connected to the first. In particular, a well-ordered society is stable because its conception of justice is stable.

In Political Liberalism, Rawls continues to speak of the stability of a conception of justice (142) or of a well-ordered society (38). However, at certain points in the text, he seems concerned that his terminology is confusing those who think of stability as a property of a political order, considered as an objective set of institutions. At these points, he adopts this ordinary meaning, and then he takes pains to add a qualification, saying that he is concerned with stability "for the right reasons" (xli). For instance, he writes that, "As always, stability means stability for the right reasons. This implies that the reasons from which citizens act include those given by the account of justice they affirm—in this case the comprehensive doctrine of justice as fairness [6]—which characterizes their effective sense of justice" (PL xlii). In Lecture IV of Political Liberalism, "The Idea of an Overlapping Consensus," he writes: "what counts is the kind of stability, the nature of the forces that secure it . . . Citizens act willingly so as to give one another justice over time. Stability is secured by sufficient motivation of the appropriate kind acquired under just institutions. [9] (PL 142-3). Again, this new formulation—stability for the right reasons—is logically connected to the other definitions of stability. For instance, a just society is stable for the right reasons if and only if it is a (stable) well-ordered society."¹⁸

Rawls's various definitions and discussions of stability are united around a single substantive preoccupation. On the one hand, Rawls wants the correct principles of justice to be realized in the basic structure of society. On the other hand, he wants the unanimity condition to be met. By characterizing certain conceptions of justice and societies as stable, he points to the fact that some are not. Stability is desirable rather than inevitable, and thus an object of concern. Fundamentally, Rawls is concerned with the ability of a certain kind of society to perpetuate itself over time. Now, according to this chapter's reading of Rawls, justice is required for the realization of objective freedom. Moreover, subjective freedom is realized if and only if the unanimity condition is met. Finally, Rawls's concern with stability is a concern that a just society nurture the unanimous endorsement of the principles of justice over time. Thus, in confronting the question of stability, Rawls is concerned with the successful self-perpetuation of a free society.¹⁹

This statement should be qualified by the observation that Rawls's interest in unanimity condition is theoretically complex. In "A More Democratic Liberalism," Joshua Cohen makes the case that, for Rawls, "a moral consensus on political fundamentals is a basic good for at least four reasons" (1516). First, Cohen says, "the existence of a moral consensus on [a conception] increases the likelihood that social order will stably conform to the conception [68]" (1516). Here, Cohen uses the term "stability" in the ordinary sense. He points to the fact that order is one concern of Rawls, though not the concern that distinguishes him as a political theorist. Next, Cohen states, "a moral consensus promotes a variety of specific values of considerable importance" (1516). If citizens are assumed to act on the principles they share, "consensus . . . increases social trust and harmony, supports social peace, simplifies decision making,

[and] reduces monitoring and enforcement costs by encouraging a willingness to cooperate.”²⁰ Third, Cohen asserts, citizens who do not choose the principles of justice may nevertheless be “politically independent and self-governing” (1516) if they endorse the principles. In his footnote to this discussion, Cohen refers to Rawls’s definition of full autonomy. Translated into the terminology of my dissertation, Cohen points here to the fact that a society that sustains a consensus is one whose citizens are subjectively free. Finally, Cohen says, “Under conditions of political consensus, citizens achieve a form of mutual respect [because] each offers as reasons for a decision only considerations that others who are subject to political power take as reasons, and state power is exercised only within the bounds of these reasons” (1517). The last case assumes that consensus not only exists but that everyone knows the other members of society endorse the same principles of justice.

Cohen’s discussion shows that, for Rawls, subjective freedom is among the values promoted by a consensus on the principles of justice, but it is not the only such value. The unanimity condition is important for a number of reasons. Recall that Rawls’s concern with stability is the concern that a just society nurtures the unanimous endorsement of its governing principles. A complete interpretation of Rawls’s concern with stability would require discussing all of the values that he thinks the unanimity condition secures or promotes. My dissertation singles out the value of subjective freedom, but it does not deny that Rawls cares about unanimity for various other reasons.

Keeping that caveat in mind, it should now be clear that, by treating stability as a concern, Rawls calls into question the persistence of the fully free social world. However, analyzing the concept of stability does not explain *why* Rawls thinks that the

persistence of a fully free social world will be a problem. The immanent critique offers one answer: the fully free social world will not persist because objective freedom will undermine subjective freedom. The dissertation will argue that Rawls's writings in fact suggest two different mechanisms by which objective freedom might subvert subjective freedom. I call the corresponding problems that arise the Public-Private Problem and the Intelligibility Problem. I am now in a position to sketch the logic behind the two mechanisms and note the relevant sources for the reconstruction of each problem. Not surprisingly, one important source is Rawls's discussion of stability in A Theory of Justice.

VII. The Public-Private Problem and the Intelligibility Problem

The dissertation's introduction sketched the two routes by which Rawls thinks that the objective freedom of a society may threaten its subjective freedom. This section returns to the topic, and goes into more detail. Given the way the general concepts of the subjective freedom of individuals and societies are linked, the objective freedom of a society threatens its subjective freedom by threatening the subjective freedom of its members. Rawlsian citizens are subjectively free if and only if they affirm the principles of justice, understand how they apply to the basic structure of society, and act from that understanding and affirmation. Now, Rawls has an account of how citizens will be socialized into the principles of justice as they grow up. But Rawls argues that such an account of socialization is insufficient to ensure that citizens will endorse the principles of justice. Because Rawls thinks that all citizens are reflective, each can question the results of their socialization. At this point, they may cease to affirm the principles of justice. Rawls traces this possibility to an intrinsic feature of the objectively free social

world: its differentiation. He distinguishes two main ways that the objectively free society is differentiated, each corresponding to a problem described in this section.

First, the objectively free social world is differentiated into a basic structure, regulated by the two principles of justice, and a private sphere. The explanation of how the two parts fit together—the social components and logic of each part—is based on the theoretical account offered by Rawls’s political philosophy. The derivation of the principles of justice is especially complex, relying as it does on a complicated procedure of construction.²¹ The rationale for the content of the principles is simply not obvious at first glance. The reflective citizen may not be able to make sense of the objectively free social world on his own. If he remains in a state of confusion, he may suspend his judgment and reject his social world in general. If he actively attempts to resolve his confusion through further reflection, he will be prone to a particular cognitive error, explained in the following way.

Rawls thinks that a just basic structure will allow many different conceptions of the good to flourish within the private sphere. This outcome is natural and healthy. As a result, there is no guarantee that an individual’s private conception of the good will correspond to the principles of justice. Given a vividly compelling private conception of the good and principles of justice that appear abstract and unmotivated, the reflecting citizen may well reject the principles of justice for his private conception of the good in matters involving the basic structure of society. Here the citizen’s mistake is an instance of a type of cognitive error: failing to recognize an entity as a part of a whole, and illegitimately extending the entity’s governing principles beyond its boundaries. In this case, the citizen’s private outlook, nurtured by public institutions, encroaches on these

very institutions. The upshot is that the citizen will actively reject the laws and norms of the existing basic structure.²²

From A Theory of Justice onward, Rawls is concerned by the possibility described above. In my dissertation, I will refer to the corresponding social problem as the “Public-Private Problem.” The main sources for describing this problem are the essay “Kantian Constructivism in Moral Theory” and Rawls’s writings on stability in Part III of A Theory of Justice. Chapter five of my dissertation describes the problem in full.

The dissertation explains the Public-Private Problem by the objectively free society’s differentiation into a basic structure and a private sphere. The dissertation will also show that the basic structure of an objectively free society is itself differentiated and that this feature causes a second problem. First, the dissertation will introduce a particular account of institutional functioning. According to this account, institutions achieve their goals by specifying a variety of offices, each with its own code of conduct. The goal of the institution as a whole is promoted by the interaction of its officials, many of who do not promote it directly. I will show that, throughout his career, Rawls adheres to this account of institutional functioning. Therefore, in a just basic structure, some officials will not directly promote the principles of justice. Now, this fact is not due to the substantive purpose of the basic structure. According to this account of institutions, unjust institutions would work in the same way. But it is also true, according to this account, that just—objectively free—institutions are necessarily differentiated.

The differentiation of the basic structure of society poses a threat to subjective freedom when an individual occupying a public office begins to reflect on the norms of his office. The true explanation for the norms of an office will take into account the

office's relationship to the basic structure as a whole. As a result, the explanation may be long and complex. The official may have trouble reaching this explanation through solitary reflection. Moreover, the reflective official may commit a version of the cognitive error described above. That is, he may reject the principles of justice in favor of his office's norms in all aspects of his public life. He then will reject the principles of justice and the norms governing parts of the basic structure outside of his official orbit. At this point, he will no longer be subjectively free. The dissertation calls the problem just described "the Intelligibility Problem," because it is caused by the fact that the basic structure is not immediately intelligible to its participants. Locating these arguments in Rawls's writings will require a significant reconstruction and extension of the texts at hand. Still, chapter seven of the dissertation makes the case that much of the relevant material can be extracted from Rawls 1955 paper on rule utilitarianism, "Two Concepts of Rules," and §10 of A Theory of Justice.

This section has now sketched the arguments from Rawls's texts, explicit and implicit, that motivate the immanent critique. Now I can state the essential intuition behind the common solution to the two problems in way that sounds deceptively simple. If the reflective citizens of an objectively free society tend at first to over-extend their private or official loyalties, it is nonetheless possible to check this tendency. A theoretical account exists that explains how the citizen's private conception of the good and official roles fit into the social world as a whole. In brief, this theory must be made available to citizens through public, educative institutions. Then citizens can master the theory and restore their subjective freedom. The burden of the rest of the dissertation is to show that

this argument is more surprising and more complex than it may appear here, and that Rawls borrows from Hegel's account of reconciliation to make it.

VIII. Conclusion

This chapter has reconstructed Rawls's conceptions of subjective and objective freedom from his body of work. It contends that Rawls's notions of autonomy and of the well-ordered society are the keys to connecting his writing on justice to a polity's two dimensions of freedom. The chapter then contrasts Rawls's conceptions of objective and subjective freedom with Hegel's from the previous chapter. While each author grounds his conceptions in an account of freedom as the self-determination of the will, I claim that particular features of each account make Hegel's conception less individualistic and less egalitarian than Rawls's. The chapter then introduces Rawls's definition of stability and argues that, by treating stability as a problem, Rawls calls into question the reproduction of a fully free social world. Finally, this chapter sketches two paths by which Rawls thinks objective freedom threatens subjective freedom, which I name the Public-Private Problem and the Intelligibility Problem. These two problems constitute Rawls's presentation of the immanent critique. The dissertation claims that, in both cases, Rawls's working out of the problem and his statement of its solution are influenced by his reading of Hegel. The next chapter begins the substantive argument for that claim.

Works Cited

Barry, Brian. "John Rawls and the Search for Stability." Ethics 105 (1995): 874-915.

Beck Hoy, Joyce. "Hegel's Critique of Rawls." Clio 10 (1981): 407-22.

Bentham, Jeremy. An Introduction to the Principles of Morals and Legislation. Ed. J. H. Burns and H. L. A. Hart. London: Methuen, 1970.

Cohen, Joshua. "A More Democratic Liberalism." University of Michigan Law Review

- 92 (1994): 1503-46.
- Hart, H.L.A. "Rawls on Liberty and its Priority." The University of Chicago Law Review 40 (1973): 534-555.
- Houlgate, Stephen. "Hegel, Rawls, and the Rational State." Beyond Liberalism and Communitarianism: Studies in Hegel's Philosophy of Right. Ed. Robert R. Williams. Albany, NY: SUNY Press, 2001. 249 - 73.
- Mill, James. James Mill: Political Writings. Ed. Terence Ball. Cambridge: Cambridge UP, 1992.
- Mill, John Stuart. On Liberty and Other Writings. Ed. John Gray. Oxford: Oxford UP, 1998.
- Neuhouser, Frederick. Foundations of Hegel's Social Theory: Actualizing Freedom. Cambridge, MA: Harvard UP, 2000.
- Rawls, John. "The Basic Liberties and Their Priority." Liberty, Equality, and the Law: Selected Tanner Lectures on Moral Philosophy. Ed. Sterling M. McMurrin. Salt Lake City, UT: University of Utah Press, 1987. 1-87.
- . Collected Papers. Ed. Samuel Freeman. Cambridge, MA: Harvard UP, 1999.
- . Justice as Fairness: A Restatement. Cambridge, MA: Harvard UP, 2001.
- . Political Liberalism. New York: Columbia UP, 1996.
- . A Theory of Justice. Cambridge, MA: Harvard UP, 1971.
- Schwarzenbach, Sibyl A. "Rawls, Hegel, and Communitarianism." Political Theory 19 (1991): 539-71.
- Sensat, Julias. "Classical German Philosophy and Cohen's Critique of Rawls." European Journal of Philosophy 11 (2003): 314-53.
- Sidgwick, Henry. *The Methods of Ethics*. Indianapolis, IN: Hackett, 1981.

Notes

¹ For the definitions of the general concepts of objective and subjective freedom, see chapter one, section II, and chapter two, section I.

² Utilitarian theories of justice admit of many variations, which I omit for reasons of space.

³ Major utilitarian theorists include Jeremy Bentham, James Mill, and Henry Sidgwick. John Stuart Mill also produced an influential body of moral philosophy that most scholars consider utilitarian.

⁴ In 1999, a revised edition of A Theory of Justice appeared. In the preface for this edition, Rawls explains that the text incorporates additions and clarifications to his theory made by him through 1975 (xi). He says, however, that many of the changes are better explained in other publications that appeared in the early 1980s (xii-xiii). Throughout the dissertation, unless I explicitly note otherwise, I quote from the original 1971 version of A Theory of Justice. I generally cite the later essays, such as “Social Unity and Primary Goods,” to explain the ways Rawls elaborated his theory after 1971.

⁵ I omit discussion of the principle of just savings.

⁶ These criticisms are made in Hart’s 1973 essay, “Rawls on Liberty and Its Priority.”

⁷ Rawls’s omission in the second principle of the just savings principle is not an abandonment of it, but is intended to streamline the statement of the two main principles.

⁸ In this essay, Rawls does develop the grounds for the priority of the first principle over the second in greater detail. His argument here is supported by the moral conception of the person. I introduce and discuss this notion the next section of this chapter.

⁹ See Justice as Fairness, p. 42, for a final statement of the principles of justice.

¹⁰ Rawls thinks that model conceptions are ideals in the sense that people should actively strive to realize them. Unfortunately, Rawls does not make that point clear when he first introduces the notion of model conceptions in “Kantian Constructivism in Moral Theory” (CP 307). Other remarks in his writings bear out this interpretation of model conceptions, though. For instance, in the preface to the 2nd edition of A Theory of Justice, Rawls speaks of “a moral conception of the person that embodies a certain ideal” (xiii). In “The Basic Liberties and their Priority,” he says that “it is the full autonomy of active citizens which expresses the political ideal to be realized in the social world” (20). Or, in A Theory of Justice, §72, he writes of a well-ordered society: “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. A perfectly just society should be part of an ideal that rational human beings could desire more than anything else once they had full knowledge and experience of what is was” (TJ 477).

¹¹ One objection to this presentation is that rational autonomy is a property of parties in the original position rather than of actual people. Based on the discussion in “Kantian

Constructivism,” this objection holds some weight. However, Political Liberalism makes it clear that, while it is “artificial” (PL 72) to think of an actual agent as *only* rationally autonomous, rational autonomy is *a* property of citizens. In this work, Rawls says that “. . . citizens are rationally autonomous in two ways—they are free within the limits of political justice to pursue their (permissible) conceptions of the good; and they are motivated to secure their higher-order interests associated with their moral powers” (74).

¹² Rawls generally prefers to confine the term “rational” to means-end rationality (CP 316). The dissertation tends to use “rational” and “reasonable” interchangeably.

¹³ In particular, chapter five of the dissertation discusses Brian Barry’s 1995 review article, “John Rawls and the Search for Stability.” I do not agree with all that this article contains, but I do endorse its claim that a concern with stability drives the changes in Rawls’s political theory.

¹⁴ Publicity has, according to Rawls, three levels (CP 324-5), and the availability of the full justification of the two principles is the final, third, level.

¹⁵ Section VII of this chapter considers at greater length Cohen’s treatment of the subjective freedom of individuals in Rawls’s political philosophy. In footnote 26 to that section, I discuss some other ways that the Cohen article overlaps with my dissertation’s argument.

¹⁶ Admittedly, Hoy’s piece predates “Kantian Constructivism,” which distinguishes explicitly between rational and full autonomy. However, given that the definition of “autonomy” in TJ is *full* autonomy, her characterization of Rawls is still problematic.

¹⁷ Clearly, not all contrasts between the two authors range along these dimensions. For instance, Rawls has no official account of human needs: the list of primary goods exhausts the objects of distribution in Rawls’s theory of justice. Although Rawls recognizes that, in the real world, citizens will use part of their income and wealth to meet their needs as human beings, Rawls’s theory does not treat human needs directly: the conception of the Rawlsian citizen is not one of a bearer of physical and material needs. Instead, income and wealth are characterized more generally as the means for realizing one’s conception of the good. Hegel, on the other hand, does work with an account of human needs, as discussed in chapter two of this dissertation.

¹⁸ This terminological innovation is not the major change in Rawls’s theory represented by Political Liberalism. The change does not consist of the definition of stability, but instead of the mechanisms that ensure stability.

¹⁹ Here I am arguing that Rawls’s interest in stability derives from a concern with the ability of free society’s parts to work together to ensure its persistence as a whole over time. Political theorists in the Hegelian tradition might instead speak of a free society’s

ability to “reproduce” itself. I conjecture that Rawls is aware of this terminology, but avoids it because of the analogy it makes between the state and biological entities.

²⁰ He adds that “if public debate and decisions reflect the consensus—[it] reduces alienation from public choices because citizens embrace the norms and ideals that guide those choice” (1516). However, the absence of alienation can be interpreted as the presence of subjective freedom, so I do not think this value is really distinct.

²¹ In general, the use of a social contract-style argument as part (or all) of an account of the objective freedom of citizens does not guarantee that the subjective freedom of citizens follows automatically. One agreement is a hypothetical and non-historical device, while the other agreement is a desirable feature of an empirical social world.

²² Cohen also says that Hegel and Rawls are both concerned with stability (1505-9), and that Rawls’s writing contains an internal critique based on freedom (1517). However, his emphases are different. First, he does not focus on stability as a concern with subjective freedom, but rather with social unity in general. Second, he says that “expressive and associative liberties” (1517) are necessary because Rawls thinks that in a free society the endorsement of the principles of justice must be voluntary rather than coerced. However, he does not mention that associative and other freedoms also are designed to enable members of society to frame, revise, and rationally pursue their conceptions of the good.