

## Exploring the Theme of Reflective Stability: John Rawls' Hegelian Reading of David Hume

Margaret Meek Lange  
Columbia University

**Abstract.** Selections from John Rawls' writings on historical figures were published in the 2000 *Lectures on the History of Moral Philosophy*. My paper discusses Rawls' treatment of Hegel and David Hume. It focuses on the following themes: the average individual's understanding of his social institutions, the psychological mechanism of "reflection" as a source of change in that individual's understanding, and the role of individual reflection in guiding social reproduction and change. I argue that these questions are central concerns of Hegel's idealist philosophy; that Hegel's position is nuanced; and Rawls recognizes both the centrality and the subtlety of Hegel's discussions. Next, I show how Rawls attends to these themes in Hume's moral philosophy as well. Since these themes are less obvious features of Hume's thought, I argue that Rawls performs a Hegelian reading of Hume. I close with a discussion of these writings' relevance to scholarship on Rawls' own work.

**Key words:** Hegel, Hume, institutions, rationality, Rawls, reflection.

During his life, John Rawls was well known as the most prominent liberal political theorist writing at the time. Because of his work as a teacher, he also had an informal reputation as a reader of historical sources, despite not much publication on these topics. Selections from John Rawls' writings on historical figures were published in the 2000 *Lectures on the History of Moral Philosophy*. This work includes sections on Hume, Leibniz, Kant, and Hegel.<sup>1</sup>

Rawls' originality as a reader makes the *Lectures* a rich resource for scholars of the four historical figures that he treats. The text also sheds light on Rawls himself as a thinker. In the *Lectures*, Rawls tracks the same topics across the works of multiple theorists. The repetition of topics makes Rawls' intellectual preoccupations very obvious. Identifying Rawls' theoretical concerns in the *Lectures* can in turn inform the study of his own political theory.

This paper will use the *Lectures* to reveal Rawls' interest in a set of interlocking themes: the average individual's understanding of his social institutions, the psychological mechanism of "reflection" as a source of change in that individual's understanding, and the role of individual reflection in explaining social reproduction and change. These themes are also central concerns of Hegel's political theory. The first part of my paper describes Rawls' accurate presentation, in the *Lectures*, of this material in Hegel's philosophy. Next, I turn to the Hume section of the *Lectures*. I propose that Rawls performs a "Hegelian reading of Hume," since he seeks to reconstruct Hume's less obvious position on these themes. Finally, at the end of my paper, I discuss how Rawls' readings of these

---

1] I presented an earlier version of this paper at the 2008 annual meeting of the American Political Science Association. I thank my discussant, Michael Frazer, for his comments.

two historical sources can shed light on his own political theory. I argue that Rawls' work on reflection and social stability takes elements from both Hume and Hegel, and that he seeks to claim Hume as a predecessor within the Anglo-American tradition.

### I. RAWLS' READING OF HEGEL

This section of my paper explicates Rawls' reading of Hegel in the last 42 pages of the *Lectures*. The first text that Rawls treats in these pages is Hegel's major mature work of political philosophy, the *Grundlinien der Philosophie des Rechts*, or *Elements of the Philosophy of Right* ([1821] 1991). Rawls also occasionally turns to a second text for his exposition, the *Introduction to The Philosophy of History* ([1820s] 1988). This short work combines a summary of the main content of the *Philosophy of Right* with a summary of Hegel's philosophy of history.

Although Rawls' reading of these texts is at times controversial, more often it is fairly conventional. Still, Hegel's writing is difficult and his philosophical doctrines require careful elaboration. This section will first summarize Hegel's own view of the average individual's capacity for rationality, both practical and theoretical. I do so by referring to Rawls' text, thus showing that he correctly presents these aspects of Hegel's thought. I consider certain features of Hegel's view of individual human agency. Then the section turns to Hegel's notion of reconciliation. Applied to the social world, reconciliation is an affirmative philosophical attitude of an individual toward his social institutions. I show that Rawls recognizes and underscores the importance of reconciliation to Hegel. Next, I claim that, while Hegel's typical agent lacks full practical rationality, the reconciled individual meets a demanding standard of theoretical rationality. I end the section with a consideration of the connections between practical and theoretical rationality. For Hegel, I ask, when can changes in individual theoretical rationality guide individual action and thereby shape the social world as a whole? I argue that Hegel's answer here is not simple, and that Rawls correctly acknowledges its complexity.

I divide the topic of agency into two parts, the agency of individuals before the modern era, and the agency of the members of the modern social world. To explore the first question, it is necessary to turn to Hegel's philosophy of history, and Rawls' consideration of it. This topic is not the focus of the *Philosophy of Right*. Still, it is treated at the end of this text (Hegel 1991, §§341-360). Moreover, the topic is central to the *Introduction to The Philosophy of History*, and Rawls occasionally cites this work. According to Hegel, human history is essentially progressive. Its goal is often described as the self-actualization of *Geist*, or Spirit. Now, exploring the meaning of this phrase—the role of Spirit in Hegel's philosophy—is beyond the scope of this paper. For my purposes, it is sufficient to establish that the self-actualization of Spirit requires the eventual emergence of the modern social world. The *Philosophy of Right* contains detailed descriptions of three major institutions: the family (Hegel 1991, §§158-181), civil society (§§182-256), and the political state (§§257-360). Hegel regards these institutions of "Ethical Life," or *Sittlichkeit* as a descrip-

tion of the rational content immanent in the modern world at the time of his writing.<sup>2</sup> That is, the modern social world is essentially composed of the institutions of *Sittlichkeit*. These are defined by their ability to realize multiple forms of freedom within one stable, self-reproducing social whole.

Since Hegel views history as progressive, earlier events can be understood in terms of their contribution to later ones. Certain key actors can also be understood in terms of their contribution to history's progressive movement. However, Hegel does not think that world-historical actors understand themselves the way we, as observers of history, later come to see them. Rawls makes this point more than one his lectures, citing Hegel's famous phrase, "the cunning of reason" (Hegel 1988, 35):

Hegel often characterizes the greatness of great historical figures in terms of their contribution to the progressive development of the institutional structure of human social life. The actions of historical agents over time unintentionally realize great social transformations that philosophy, looking back after the fact, understands in terms of the cunning of reason. Great figures seek their own narrow ends, yet unknown to them they serve the realization of *Geist*. (Rawls 2000, 369)

The "cunning" of reason references Hegel's view of reason as a superhuman agent, which directs the course of world history through the actions of human individuals. A bit earlier Rawls summarizes Hegel's example of Martin Luther. With Luther, the Protestant Reformation began, which in turn led to the "division of Christendom" (Rawls 2000, 347). Eventually, this situation led to the establishment of freedom of religion as a solution to religious strife, a freedom that was later affirmed, not as a *modus vivendi* only, but as good for its own sake (Rawls 2000, 347-8). Therefore, although Luther certainly did not intentionally promote freedom of religion, his life work indirectly served to establish it. Rawls writes:

Ironically, Martin Luther, one of the most intolerant of men, turns out to be an agent of modern liberty. This is an aspect of history that Hegel emphasized—that great men who had enormous effects on major events of history usually never understood the real significance of what they had done. It is as if they are used by a providential plan unfolding through time and embedded in the flow of events. (2000, 348)

These comments show that not all world-historical individuals understand the character of their actions.

Now, I consider a second kind of agent, the agent existing in the modern social world—the social world defined by the institutions of *Sittlichkeit*—whose actions are necessary to reproduce those institutions. One of Hegel's assumptions is that many members of modern *Sittlichkeit* will obey the laws and norms of their society without fully grasping their philosophical significance. Rawls writes: "Hegel wants us to find our moral compass in the institutions and customs of our social world itself, as these institutions and customs

---

2] Following usage among contemporary writers on Hegel such as Frederick Neuhouser and Alan Patten, I retain the term *Sittlichkeit* in the original German.

have been made part of us as we grow up into them and develop habits of thought and action accordingly” (2000, 333). In the *Philosophy of Right*, *Sittlichkeit* is described as providing us with a “second nature” (Hegel 1991, §151).

Thus, Hegel has philosophy of history that posits that world-historical individuals change the structure of social institutions without intending to do so. He also thinks that the individual actions that reproduce the structure of the modern social world are often done from custom or habit. Neither type of agent meets the most stringent standard of practical rationality that could be formulated. Because the true significance of their actions is not accessible to them, Hegel’s world-historical agents lack full practical rationality. The same is true for the average member of *Sittlichkeit*, who endorses his institutions but does not understand exactly how they work to provide for his freedoms.

Now I turn to Hegel’s notion of reconciliation. Rawls opens his lectures on Hegel by writing, “I begin by noting Hegel’s view of philosophy as reconciliation” (2000, 31). Rawls continues: “Hegel thinks that the most appropriate scheme of institutions for the expression of freedom already exists. It stands before our eyes” (331). Recall that Hegel believes that modern social world is essentially composed of the three major institutions of *Sittlichkeit*. Moreover, Hegel thinks that the essence of the human individuals is to be free, and the institutions of *Sittlichkeit* enable human freedom. Therefore, Hegel thinks that the modern social world is worthy of endorsement by rational individuals.

However, within the modern social world are individuals who both seek a justification of modern institutions and who do not yet understand that the scheme of institutions realize freedom. Sometimes Hegel refers such individuals as alienated (Inwood 1992, 35-8). Rawls explains Hegel’s solution to this problem: “The task of philosophy, especially political philosophy, is to comprehend this scheme in thought. And once we do this, Hegel thinks, we will become reconciled to our social world” (2000, 331). Hegel’s political philosophy can demonstrate to these individuals the rationality of their social world, by describing to them the institutions of *Sittlichkeit*. Rawls notes: “Philosophy in this role is not merely an academic exercise. It tells us something about ourselves; it shows us our freedom of the will—that we have it through institutions, not in other ways” (331). Reconciliation—in German, *Versöhnung*—is the attitude achieved by individuals to their social world once they are philosophically enlightened. As Rawls emphasizes, as Hegel uses the German word, “reconciliation” does not imply resigned acceptance (331), but rather full affirmation.

I now consider two further aspects of reconciliation. First of all, for Hegel, achieving reconciliation requires grasping the whole social world in its essence—not fixating on any one particular part of it in isolation from the remainder. The importance of the whole is basic to Hegel’s thought in general. Hegel thinks that reality (*Wirklichkeit*) forms a system, a hierarchically organized entity whose pieces are defined in terms of their place in the overall organization (Inwood 1992, 266). Thus, the entirety of reality must be grasped in order to have a full understanding of any piece of it. While the whole is the proper object

of philosophical thought, concentration on one part of reality to the exclusion of the rest is a major source of philosophical error. Now, towards the end of the *Lectures*, Rawls writes:

What raises human life above the workaday *bürgerliche* world is the recognition of the universal interest of all citizens in participating in and maintaining the whole system of political and social institutions of the modern state that make their freedom possible. Citizens knowingly and willingly acknowledge this universal (collective) interest as their own, and they give it the highest priority. They are ready to act for it as their ultimate end. This is the goal of the project of reconciliation. (2000, 355)

This passage shows that Rawls recognizes that the entirety of the modern social world, inasmuch as it realizes the institutions of *Sittlichkeit*, is the object of reconciliation.

Also important to underscore is the connection between reconciliation and reflection. As stated above, Hegel thinks that reality forms a system, and must be comprehended in its totality to be comprehended truly. The philosophically enlightened subject's knowledge of reality must be systematically organized to mirror the reality it represents (Inwood 1992, 266). However, Hegel thinks that the subject has to acquire this knowledge, and here is where the notion of reflection comes in. Roughly speaking, reflection is a mental operation performed by a subject to render his or her beliefs more systematic.

An initial act of reflection, Hegel thinks, is insufficient to produce a fully systematic account of reality. Hegel refers to the "abstract" nature of concepts that treat the parts of reality in isolation from their context in the whole (Inwood 1992, 30-31). Reflective moves first lead into philosophical error by producing abstractions. Sometimes Hegel speaks of the limited nature of the "philosophy of reflection" (*Reflexionsphilosophie*) (Inwood 1992, 249), which has a tendency to view the world dualistically because it does not understand how the parts it separates are related to each other through the whole.

In the Preface to the *Philosophy of Right*, Hegel says that modernity encourages reflection, and that this trend is a positive development: "It is a great obstinacy, the kind of obstinacy which does honour to human beings, that they are unwilling to acknowledge in their attitudes anything which has not been justified by thought—and this obstinacy is the characteristic property of the modern age, as well as being the distinctive principle of Protestantism" (1991, para. 14). However, the increased tendency toward reflection has produced alienated individuals, who cannot see that their social world promotes their freedom because they are limited by the "fetter of some abstraction or other" (Hegel 1991, para. 13). Now, according to Hegel, the solution to reflection-induced alienation is not a return to a pre-reflective state, but further reflection. Later reflective moves correct the errors of the previous ones and produce an accurate body of systematic knowledge, leading to reconciliation through philosophy for the individual. Hegel summarizes: "It has become a famous saying that 'a half-philosophy leads away from God' . . . 'whereas true philosophy leads to God'; the same applies to philosophy and the state" (1991, para. 14). This claim does not mean that the earlier stages of reflection are unnecessary; they are part of reflective thinking's coming to a proper understanding of reality. In Hegel's phi-

losophy in general and the Preface to the *Philosophy of Right* in particular, reflection is first the cause of alienation and then the cure for it.

In the *Lectures*, Rawls recognizes the basic point that, for Hegel, reflection ultimately facilitates reconciliation. He writes: “[W]hen in our reflections we understand our social world as expressing our freedom and enabling us to achieve it as we live our daily life, we become reconciled to it” (Rawls 2000, 332). Elsewhere he says that Hegel thinks individuals should “belong to a rational (reasonable) social world that [they] on reflection can accept and be reconciled to as meeting their fundamental needs” (Rawls 2000, 333). In other passages, Rawls recognizes that reflection can produce alienation in individuals as well as reconciliation (2000, 345-6).

Now I maintain that the reconciled individual meets a demanding standard of theoretical rationality. I take it that the goal of the theoretically rational subject is to acquire knowledge: justified true belief. The reconciled individual bases his affirmation of his social world on knowledge of its workings. That is, he has a complete, true understanding of it based on political philosophy. Therefore he is theoretically rational.

I now consider the possible connections between theoretical and practical rationality in Hegel’s political philosophy, and between practical rationality and the shape of the social world. First of all, I note that reconciliation cannot guide structural change since it always will come after structural change. So far we have seen that Hegel thinks many world-historical actors, such as Luther, did not know the historical meaning of their actions. In the Preface to the *Philosophy of Right*, Hegel makes an even stronger claim. He writes: “A further word on the subject of issuing instructions on how the world ought to be: philosophy, at any rate, always comes too late to perform this function. As the thought of the world, it appears only at a time when actuality has gone through its formative process and attained its completed state” (Hegel 1991, para. 15). Here Hegel says that the proper philosophical understanding of principles awaits the existence of a society that realizes them. Therefore, principles underlying social institutions never successfully guide individuals in making fundamental social change.

However—and this is the second possibility—Hegel does think that theoretical reflection can upset a given social structure by affecting the agency of its members. In fact, he thinks that this occurred in the world of ancient Greece (1991, §185A). Rawls summarizes: “This unreflective form of life inevitably becomes unstable and falls into decay upon the appearance of reflective thought” (2000, 345). Once the individuals no longer theoretically affirm their social structure, they will not act to reproduce it, and it will pass away.

Finally, I consider a third possibility. If the social structure is fully rational, the knowledge won through philosophy can then guide an individual’s compliance with it. The modern social world is an example of a fully rational social world, and thus embodies the third possibility. In fact, it is important that gains in theoretical insight can subsequently guide rational action, because one purpose of reconciliation is to ensure the individual citizen’s practical freedom. In order to be fully free within his social world, Hegel thinks that an individual must minimally endorse the laws he obeys. Now, as Rawls notes, the individual

does not have to have a full philosophical understanding of these laws and norms: “Hegel wants to show that people can and do act freely when conducting themselves on the basis of habit and custom” (2000, 333). However, once the individual is alienated, he no longer affirms the laws and norms at all, and is therefore no longer fully free. Reconciliation through philosophy shows the individual why the institutions are rational so that he can then reaffirm them. The theoretical account restores the individual’s complete practical freedom by reinstating freedom’s essential subjective dimension. Hegel makes this connection between reconciliation and freedom quite clear in the Preface. He writes:

To recognize reason as the rose in the cross of the present and thereby to delight in the present—this rational insight is the reconciliation with actuality which philosophy grants to those who have received the inner call to comprehend, to preserve their subjective freedom in the realm of the substantial. (Hegel 1991, para. 13).

At other points in the *Philosophy of Right*, Hegel shows that he thinks of subjective freedom as practical (1991, §258A). Thus, Hegel’s position that many members of the social world lack full practical rationality does not mean that some are not fully practically rational. Furthermore, those few achieve their practical rationality through theoretical reflection.

Moreover, according to Hegel’s metaphysics, not only individuals but also larger social entities can possess freedom as a property. However, the freedom of a social entity and the freedom of the individuals that compose it are related for Hegel. A proper philosophical understanding of the social world by its philosophically minded members is necessary for its freedom as well as their own (Hegel 1991, §145). Rawls underscores this point. He writes: “[The] understanding [of individuals] makes a form of life real. The explanation is that a form of life is not fully made real or actual (*wirklich*) until it is made self-conscious” (Rawls 2000, 332). A bit later, Rawls says, “[W]hile rational social institutions are the necessary background for freedom and for individuals’ real autonomy, the reflection, judgment, and rational (reasonable) conduct of individuals are necessary to bring about the substantiality and freedom of their social world” (2000, 334). Note that “rational (reasonable) conduct” is included on this list. Now, elaborating the metaphysical assumptions supporting these claims is beyond the scope of the paper. What is important is that the practical freedom of individuals is necessary to ensure the freedom of the whole social world—and that the possibility practical freedom is ensured by the possibility of reconciliation.

Rawls refers to a social structure that can survive the reflection of its individual members as “a stable form of reflective social life” (2000, 346). Here Rawls points to the causal chain that Hegel thinks runs from individual reflection through individual action to the shape of the social world that the individual inhabits. Furthermore, these phrases in the Hegel section of the *Lectures* that join “reflection” and “stability” are worth noting for a second reason. We will see that such a construction also appears in Rawls’ writing on Hume. I now turn to that section of the *Lectures*.

## II. RAWLS' HEGELIAN READING OF HUME

Rawls' discussion of Hume is quite a bit longer than his reading of Hegel. In the first 102 pages of the *Lectures*, Rawls explicates Hume's moral philosophy, as it appears in two of his major philosophical works. These works are his *Treatise of Human Nature* ([1739-40] 1978), especially Book III, "Of Morals"; and his later *Enquiry Concerning the Principles of Morals* ([1751] 1998). Rawls devotes most of his space to the *Treatise*, and since I am following his interpretation, I will do likewise.

Hume, Rawls says, provides a psychological explanation of instrumentally rational behavior. However, he argues, Hume has no notion of "practical reason" (Rawls 2000, 37). Practical reasoning, Rawls says, is "deliberation regulated by (ostensibly) correct or valid judgments and moved by principle-dependent desires associated with rational principles," (2000, 50). In the case of a principle-dependent desire, Rawls says, "the aim of the desire, or the deliberative, intellectual activity in which we desire to engage, cannot be described without using the principles, rational or reasonable as the case may be, that enter into that activity" (2000, 47). To say Hume has no notion of practical reason is not the same as saying his view of human deliberation "is simple and not complex" (Rawls 2000, 50); indeed "it is very complex" (50). Rawls shows that Hume thinks human beings possess psychological mechanisms that allow them to choose the correct means to their ends, and to revise and organize their ends (2000, 38-43).

Next, Rawls shows that Hume also argues that moral judgments are not based on reason either. The account of moral judgment is also "psychological" (Rawls 2000, 96) inasmuch as it is based on the hypothesis that our moral distinctions are due to a moral "sense" (Hume 1978, 470). In the *Treatise*, Hume develops his argument for the existence of moral sense in Part I, sections I and II of Book III (1978, 455-471); and Rawls in turn covers this material in Lecture IV of the Hume section of the *Lectures* (2000, 78-83).

I now summarize Hume's presentation of the moral sense. This presentation includes the criteria by which it judges and the psychological mechanisms by which it operates. The object of the moral sense is the motivations of agents. Certain natural motivations, such as benevolence and affection toward one's children, are judged virtuous. Moreover, Hume notes that the motivation to act justly is also judged virtuous. By examining the objects of the moral sense in the context of social life, we see that the sense approves of motivations that guide action useful to oneself and to others. These criteria are made especially in Sections one through four of the *Enquiry* (Hume 1998). Now, the moral sense does not operate by consciously applying its criteria to its objects; that is why it is called a "sense." However, knowing what we do about human psychology, we can construct an explanation of how the sense works. In the *Treatise* this account appears in Book III, Section III, chapter I, and draws on the human propensity for sympathy and the capacity to take up a general point of view. Rawls explicates this material in Lecture V of the Hume section.

Thus, Rawls claims Hume offers separate, psychological accounts of moral motivation and of moral judgment. Rawls writes: "The virtues and vices are known to us in virtue of the peculiar moral sentiments we experience . . . What moves us to act in accordance

with our moral sentiments is a separate question altogether . . . The basis of these motives requires its own separate account” (2000, 99). In general, Rawls says that “The *Treatise* is an account . . . of morality psychologized” (2000, 85). However, later in his writings Hume, Rawls says:

A contemporary reader is likely to say: Hume’s account is purely psychological; it describes the role of morals in society and how it arises from the basic propensities of our nature. This is psychology, we say, and not moral philosophy. Hume simply fails to address the fundamental philosophical question, the question of the correct normative content of right and justice. To say this, I believe, is seriously to misunderstand Hume. (2000, 98)

This comment is a bit puzzling, given what has gone before. I now consider why Rawls makes that claim, despite his previous claims and general treatment of Hume. The answer is found in the passages in which he implicitly or explicitly refers to Hume’s view as a “fideism of nature.”

Fideism simpliciter “holds that reason is unnecessary and inappropriate for the exercise and justification of religious belief” (Amesbury 2008). By “fideism of nature,” Rawls means that Hume thinks we can rely on our native psychological propensities to guide our judgment of the world and action within it (Rawls 2000, 23). However, Hume does think that it is philosophical reflection that establishes this fact.<sup>3</sup> Now, Rawls invokes the notion that Hume’s philosophy is a “fideism of nature” in at least two contexts. One context is explaining our approval of the institutions of justice. Our moral sense approves the institutions of justice. Correct philosophy can explain why we are responding in this way; what criteria we use when we find these institutions worthy of approbation. The second use of the phrase “fideism of nature” occurs when Rawls considers Hume’s writing on our attitude toward the moral sense itself. According to Rawls, Hume thinks we will affirm the moral sense when we reflect on it.

I turn first to Hume’s writing on the institutions of justice. In order to explain Rawls’ interpretation here, it is necessary first to explain what Hume takes as these institutions. Hume thinks that the institutions of justice are composed of rules, or “conventions” (Hume 1978, 494) regulating the origin and transfer of property. He thinks that these rules can exist even in the absence of a central governing authority (Hume 1978, 539); and he thinks that everyone is made better off by their existence. To use contemporary terminology, they are a Pareto-improvement over a rule-less situation. The rules are first instituted from enlightened self-interest on the part of all parties. However, they are later morally approved of by the moral sense because the rules are to everyone’s advantage (Rawls 2000, 499). Rawls covers these basic points in Lecture III, “Justice as an Artificial Virtue.”

---

3] Given that Rawls says that, for Hume, theoretical justification of the social world occurs through philosophy, it may be that Rawls uses “fideism” in an unusual way. For more information on varieties of fideism, see the entry on fideism in the *Stanford Encyclopedia of Philosophy*.

Rawls begins by saying that “Hume’s account of justice . . . is central to his fideism of nature: he wants to show that morality and our practice of it are the expression of our nature, given our place in the world and our dependence on society” (2000, 51). Hume, Rawls notes, develops the notion of the circumstances of justice, showing that the rules of justice exist because of moderate scarcity and limited altruism (58-9). Given this background, our lives are all improved by rules for property and its transfer. In fact, Rawls says, our rules are the best we can do. Rawls writes:

If we take Hume literally here (and why not?), he is suggesting that the convention, or scheme of conventions, he is describing is the best scheme: “there is no better way to consult both these interests,” our and others’. He doesn’t mean that it is the best scheme we can imagine, much less the best scheme allowing that human beings and our situation in nature might have been different. He means it is practically the best scheme, accepting ourselves and our situation in nature as it is, without weeping and lament. (Rawls 2000, 60)

Rawls also emphasizes that “what Hume considers to be just is the whole plan or scheme” (2000, 64). Rawls notes that Hume thinks that “just acts, taken alone, are not infrequently detrimental to society” (2000, 60). However, Hume says (and Rawls quotes): “Tis impossible to separate the good from the ill. Property must be stable, and must be fix’d by general rules. Tho’ in one instance the public be a sufferer, this momentary ill is amply compensated by the steady prosecution of the rule” (Hume 1978, 497). Glossing this and related passages, Rawls writes:

the laws of property cannot determine possession and transfer according to who is best qualified at this or that moment to use this or that piece of property, as the particular utilities of the case might decide it. This is a recipe for endless disorder and quarrels, and calls forth the partialities of the natural affections, which the rules of justice are designed to restrain. (2000, 65)

In brief, the argument is this. In order to promote utility, we need a system of rules. But no workable system of rules is fine-grained enough promote utility in every instance. Therefore, if we want to understand and evaluate conventions of justice, it is a mistake to focus too closely on any particular instantiation of any rule. We must see how the system works over time as a whole.

The next consideration of the fideism of nature appears at the end of the Hume section of the *Lectures*, when Rawls turns to the very last chapter of the *Treatise*. Rawls argues that this chapter shows Hume reflecting on the moral sense itself. Once we see that the moral sense springs from our human capacity for sympathy, which we regard as a positive aspect of our nature, we are glad to endorse its responses to the world. Rawls writes:

Hume is saying that his science of human nature . . . shows that our moral sense is reflectively stable: that is, that when we understand the basis of our moral sense—how it is connected with sympathy and the propensities of our nature, and the rest—we confirm it as derived from a noble and generous source. This self-understanding roots our moral sense more solidly and discloses to us the happiness and dignity of virtue (T: 620). (2000, 100)

He adds, “this is all part of what I have called his fideism of nature” (100). Rawls sums up Hume’s general attitude: “He is utterly without lament or sense of loss, with no trace of romantic anguish or self-pity. He doesn’t complain against the world, a world that is for him without the God of religion, and the better for it” (100).

I now summarize the material that appears under the heading of the “fideism of nature,” and draw some connections to Hegel and his concept of reconciliation. Rawls does not himself make these comparisons, but the organization of both readings and the repetition of certain phrases suggest them. Rawls shows that Hume thinks that our moral judgments, though originally based on a moral sense, can be reflected on philosophically. However, it is possible to fall into philosophical error here: the moral sense is in some ways more reliable than philosophical reflection. Here, we can make a first comparison between Hume and Hegel. Hegel also thinks that pre-reflective judgments can be more reliable than initial attempts at philosophizing, due to the tendency to make certain errors. Moreover, both philosophers emphasize one particular error: Hume, like Hegel, clearly thinks that focus on the parts at the expense of the whole is one major source of philosophical error when thinking about social institutions. Finally, both thinkers do not stop at refuting their philosophical opponents, but often diagnose them as well, showing what cognitive psychological tendencies are leading them astray.

But, provided we reason clearly, we can arrive at the right account. In correctly showing how our moral sense works, Hume also finds what criteria our moral sense is responding to. Taken properly—that is, as a whole—the institutions of justice prove not to be alien to human interests, but in fact to exist only to serve them. This finding corresponds to Hegel’s contention that our modern institutions serve our interest in our freedom. It is true that Hume focuses on utility while Hegel points to freedom; but both make the assertion that existing institutions are for human beings, not the other way around. Both Hegel and Hume affirm their own social institutions.

Finally, once we know the principles and their source, Hume thinks that we can decide whether or not to endorse our moral sense’s responses. For Hume, regardless of whether or not we endorse our moral sense, we will still have certain moral experiences. However, it happily turns out that we do endorse the responses that the moral sense generates; it is reflectively stable. Rawls’ use of this phrase in this part of his interpretation points, I think, to the fact that he has Hegel in mind while reading Hume.<sup>4</sup>

Now, in his treatment of Hegel, Rawls uses the term “reflective stability” to refer to a given social world. His arguments using this concept rely on the notion that changes in our theoretical apprehension of the social world can affect that world’s continuation

---

4] Other possibilities are that Rawls has Hume in mind when reading Hegel, or that he is simply bringing a particular concept to bear on multiple authors. However, I would argue that the notion of reflective stability—if not the actual phrase—is prominent in Hegel’s work in way that it is not in Hume’s. Moreover, since Rawls clearly read Hegel, it is likely that Hegel influenced Rawls’ emphasis on reflective stability. Michael Frazer points out that a complete discussion of Hume and Hegel would have to discuss Hegel’s reception of Hume.

over time. These changes in the understanding of the members of the social world also affect the social world's final realization as fully free. These claims in turn assume that reflectively acquired beliefs can guide action. But in his writing on Hume and reflective stability, Rawls only says that our moral sense is what is made stronger. Is there a further connection between practical agency and theoretical reflection in Hume's thought, according to Rawls? Could theoretical reflection on social institutions be a source of institutional change, according to Hume's philosophy?

The answer is that Rawls is unclear on this topic. Rawls does occasionally imply that reflection will affect action in the philosophical, Humean, subject. In his first mention of the "fideism of nature," Rawls says "this underlying outlook guides his [Hume's] life and regulates his outlook on society and the world" (2000, 24, *my italics*). Rawls also says that "the idea of the practically best scheme [of conventions of justice] plays a role in explaining why Hume thinks that we should act justly even though just acts, taken alone, are not infrequently detrimental to society" (2000, 60). In both these sentences, it sounds like enlightened beliefs about the world are guiding action. However, the "fideism of nature" might also mean that there is no need to guide our actions through enlightened beliefs because our natural responses are trustworthy anyway.

Whether reflective judgments can be translated into action depends on the details of the moral psychology Hume offers. If Hume had an account of practical reason, and a notion of principle-dependent desires, one could make a direct connection between reflective moral judgments and moral motivation. However, recall that Rawls thinks Hume lacks a notion of practical reason. Thus Rawls thinks that, for Hume, judgments will not be automatically translated into action (2000, 97). The absence of principle-dependent desires is an obstacle, but not an insurmountable one. Because of the scope of the material Rawls considers, his discussion of all the mechanisms Hume posits is limited. A full answer would require returning to Hume's text. However, it is certainly true that these explanations will be less parsimonious than an explanation that posits the human capacity to translate principles directly into action.

So far, this paper has examined the readings of Hume and Hegel offered by John Rawls' *Lectures on the History of Moral Philosophy*. To summarize, Rawls correctly shows that Hegel offers a complex account of the capacity of the individual to shape his social world consciously. Hegel thinks that the individual can only do so by first reflecting on the social world he inhabits and grasping its character. Here he must overcome the tendency to make cognitive errors. Individuals who have completed this process can then affect the character of the social world consciously. In the case of the present social world, Hegel thinks that individuals will endorse it upon reflection and consciously reproduce it. Doing so will contribute to its freedom as a whole. These features of Hegel's idealistic philosophy that Rawls discusses are central to it.

Next, I argued that Rawls examines similar themes in Hume's philosophy. In Hume's case, the individual's ability to understand the social world depends on interpreting pre-given experience while overcoming cognitive errors. Furthermore, Hume thinks

that individuals will affirm the social world of the present. Once examined closely, this part of Hume's political theory is akin to Hegel's. I have argued that Rawls seeks Hegelian themes in Hume's philosophy and reconstructs less explicit aspects of it. Then, I show that, in Hume's case, Rawls' identifies second problem. Once theoretically enlightened, the individual may still lack the psychological capacity to let his knowledge guide his behavior.

### III. CONCLUSION:

#### TRACING THE INFLUENCE OF HUME AND HEGEL ON RAWLS' THOUGHT

This final section of my paper considers what we can learn by juxtaposing Rawls' own political theory and his readings of Hume and Hegel. The central results of Rawls' political theory are the two principles of justice that he thinks social institutions should satisfy. In this section, instead of examining the principles themselves, I will consider how Rawls relates the concepts of reflection and stability to the principles. First, Rawls says that he derives the principles by reflecting on the social institutions of existing liberal democracies. Now, Rawls does not claim that existing liberal democratic institutions are just. Instead, the institutions furnish the initial material that sets the reflective process in motion. Rawls describes a method of doing political philosophy that aims at a "reflective equilibrium" (1971, 20) of beliefs about justice. Beginning with beliefs about justice that our existing institutions encourage in us, we repeatedly adjust our general and particular beliefs until we have a coherent system of beliefs all of which we endorse (20). The principles of justice are among these beliefs, and the fruit of carrying out the method.

In Rawls' theory, the high-level method of reflective equilibrium coexists with other methods for generating and testing principles of justice. Rawls includes as a check on the correctness of such principles that they be "stable." The definition of the stability of principles relies on the definition of a "sense of justice" (Rawls 1971, 569) as "an effective desire to apply and to act from the principles of justice" (Rawls 1971, 568). Principles of justice are stable when "those taking part in [just] arrangements acquire the corresponding sense of justice and desire to do their part in maintaining them" (Rawls 1971, 454). Rawls' test for the stability of principles of justice is essentially a thought experiment. We conjecture what the sociological features of a just society will be over the long run and ask whether individuals who live in that society will endorse the principles of justice. Rawls argues that, in fact, over the long run a society that embodies the two principles will be affirmed by its citizens. Sometimes Rawls speaks of such a society governed by stable principles as itself stable (1971, 457).

I will not describe in detail Rawls' treatment of stability. In fact, with the publication in the 1990s of *Political Liberalism*, this account changes in significant ways. Documenting the vicissitudes of Rawlsian stability is beyond the scope of this paper. I confine myself to noting the relevance of the theme of reflection to the test for stability, an aspect of Rawls' discussion that remains constant over time. In *A Theory of Justice*, Part III Rawls explains how individuals that grow up in a just society acquire a sense of justice. However, he assumes that adults will not accept their sense of justice unreflectively, but will want to

examine the principles. Like individuals within the real world, they reflect on their beliefs until these reach an equilibrium state. Rawls assumes that, in a society governed by the two principles, the individual will ultimately decide to “preserve his sense of justice” (1971, 576). For Rawls, a society and its governing principles are stable only if the principles survive the reflection of individuals within the society.

Next, I will argue that Rawls’ writings on reflective stability bear the traces of both Hume and Hegel. Now, given Rawls’ overt comments about his influences, it is certainly plausible that he melds the thoughts of these two writers. First, the tradition of German idealism is a major avowed influence on Rawls’ thought. Rawls always makes it clear that his work is indebted to Kant, and in the *Lectures*, he says that *A Theory of Justice* “learns much from” both Kant and Hegel (2000, 330).<sup>5</sup> Second, Rawls also thinks of his political theory as compatible with a naturalistic worldview. In the 1980 essay, “Kantian Constructivism and Moral Theory,” Rawls praises John Dewey for “adapt[ing] much that is valuable in Hegel’s idealism to a form of naturalism congenial to our culture” (1999, 304). Like Dewey, whom he clearly admires, Rawls wishes to import features of German idealism that he finds attractive, while shedding other aspects of its heavy metaphysical apparatus. At points in his theory, he explicitly combines aspects of Kant and of Hume.<sup>6</sup> So, it is not surprising that Rawls’ writings on reflection and social stability take elements from Hume and the second major German idealist, Hegel.

Like Hume and Hegel, Rawls seeks the principles of justice within his own political tradition. However, unlike Hegel and Hume, he does not think the present social world is already acceptable to reflective scrutiny. To compare Rawls to Hume and Hegel, I move to the second aspect of Rawls’ work that emphasizes reflection, the thought experiment that establishes the stability of the principles of justice. First, Rawls’ account shares with Hume’s the name of the object of reflection. Hume thinks we reflect on and affirm our moral sense and Rawls says that we reflect on and affirm our sense of justice. In both cases, the analogy made to sense perception suggests that a naturalistic explanation can be given for these psychological experiences.

On the other hand, Rawls, like Hegel, connects reflection and the value of freedom. Recall that Hegel thinks that individuals who reflect on their institutions succeed in overcoming their alienation and preserving their freedom. Rawls also thinks that the citizens of a just society can realize a particular kind of freedom that he terms full autonomy. In “Kantian Constructivism and Moral Theory,” Rawls says: “For it is by affirming the first principles [of justice]... 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” (1999, 315). Without unpacking this definition in its entirety, I note that it requires the actor to act from his sense of justice. Thus, by preserving his sense of justice through reflection, the individual actor in a Rawlsian society also preserves his

---

5] For example, see Rawls 1971, §40, “The Kantian Interpretation of Justice as Fairness.”

6] For instance, see Rawls 1971, §22, “The Circumstances of Justice,” as well as §40 of the same work.

full autonomy. Stable principles of justice are those that can be honored autonomously after reflective equilibrium has been reached.

Thus, this section has briefly shown that Rawls' account of stability overlaps with elements of Hume and Hegel. It is certainly possible that, in crafting this account, he draws directly on his readings of Hume and of Hegel. If that is true, it also explains why, in the *Lectures*, Rawls reconstructs Hume's philosophy to bring out its common concerns with Hegel's thought. Rawls sees Hume as a writer within the Anglo-American tradition who combines a naturalistic outlook with an interest in reflection and in social stability. I conjecture that Rawls attends to these aspects of Hume's writings because he seeks to import them into his own theory and to claim Hume as an intellectual predecessor.

At points in his work, Rawls selectively mines the works of German idealism, combining insights from that school with elements from the Anglo-American tradition. This paper has argued Rawls uses that strategy in incorporating the themes of reflection and social stability into his political theory. It has summarized Rawls' treatment of Hegel, his original reading of Hume, and showed how contributions from both authors appear in Rawls' own work. The paper documents a specific example of a more general truth that the *Lectures on the History of Moral Philosophy* makes clear. That is, throughout his career, John Rawls' substantive concerns co-existed with an acute consciousness of his own situation in relationship to the tradition of political theory.

margaret.meek@gmail.com

#### REFERENCES

- Amesbury, Richard. 2008. Fideism. *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition). Ed. Edward N. Zalta. <http://plato.stanford.edu/archives/win2008/entries/fideism> (accessed January 25, 2009).
- Hegel, G.W.F. 1988. *Introduction to The Philosophy of History*. Trans. Leo Rauch. Indianapolis, IN: Hackett.
- . 1991. *Elements of the Philosophy of Right*. Ed. Allen W. Wood. Trans. H.B. Nisbet. Cambridge: Cambridge UP.
- Hume, David. [1739-40] 1978. *A Treatise of Human Nature*. Ed. P.H. Nidditch and L.A. Selby-Bigge. 2nd ed. Oxford: Oxford UP.
- . 1998. *An Enquiry concerning the Principles of Morals: A Critical Edition*. Ed. Tom L. Beaucamp. Oxford: Oxford UP.
- Inwood, Michael. 1992. *A Hegel Dictionary*. Oxford: Blackwell.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard UP.
- . 1996. *Political Liberalism*. New York: Columbia UP.
- . 1999. *Collected Papers*. Ed. Samuel Freeman. Cambridge, MA: Harvard UP.
- . 2000. *Lectures on the History of Moral Philosophy*. Ed. Barbara Herman. Cambridge, MA: Harvard UP.

