

Political Confucianism and Human Rights

Daniel P. Corrigan*

Abstract

This article examines the theory of human rights developed by Tongdong Bai in his Confucian-inspired political philosophy. Partly influenced by Rawls's "political liberalism," Bai seeks to offer a "political conception" of Confucianism. However, Bai's methodological approach also deviates from Rawls's approach in certain key respects, and this has significant implications for his theory of human rights. The article begins with a comparison of Rawls's and Bai's methodological approaches. It then discusses how these competing methodologies are used by each philosopher to develop a theory of human rights and international relations. Finally, the article seeks to adjudicate these competing accounts of human rights. Notably, Bai does not follow Rawls in offering a "political conception" of human rights, one which recognizes the role of human rights in mediating international relations between states. While Rawls's political conception of human rights has been the subject of criticism, it is shown that even a revised version of this theory presents challenges for Bai's account. The article concludes by offering suggestions about how Bai's theory of human rights should be revised in order to adhere to his methodological approach. The issues raised in this article present a challenge not only for Bai, but for any attempt to develop a Confucian theory of human rights.

Keywords: Human rights, Confucianism, political philosophy, international relations, Rawls, Tongdong Bai

* Daniel P. Corrigan is Assistant Professor of Philosophy at Marymount University. E-mail: daniel.corrigan@marymount.edu

In *Against Political Equality*, Tondong Bai offers an interesting and provocative Confucian-inspired approach to political philosophy. While Bai's book makes valuable contributions to many topics and raises an array of issues for discussion, this article focuses on his theory of human rights. Bai's approach to Confucianism is partly influenced by Rawls. However, he also deviates from Rawls in certain key respects, and this has significant implications for his theory of human rights. The article begins with a comparison of the methodological approaches adopted by Bai and Rawls, then explains how these approaches are used to develop their respective theories of human rights and international relations, and finally seeks to adjudicate the differences between these theories. Ultimately, the article offers suggestions about how Bai's theory of human rights should be revised in order to consistently adhere to his methodological approach.

I. Methodological Approaches

Bai's approach to political philosophy is inspired by the early Confucians, in particular Confucius and Mencius. He argues that the Spring and Autumn and Warring States period (SAWS) from 770 BC to 221 BC, the period in which these philosophers lived, was very similar to the situation in early modern Europe. The SAWS came about with the fall of the Western Zhou feudal empire. This led to the formation of various states that share the characteristics of modern nation states: large, populous, states of strangers, with a plurality of values and under centralized control. Due to these features of the SAWS, Bai argues that the early Confucian philosophers can be understood as addressing the problems of modernity, which are the same problems confronted by early modern European philosophers many centuries later. As a result, Bai contends, Confucius and Mencius can be understood as doing political philosophy, rather than engaged in developing a moral metaphysics. According to Bai, the works of these philosophers should be "read 'ametaphysically,' as a political conception" (28). Reading the early Confucians in this way will yield a "thin" version of Confucianism, which can be "endorsed by Confucians of competing schools, and even

by people of different comprehensive doctrines" (29).

Bai's suggestion of a "thin," "political conception" of Confucianism reveals the Rawlsian influence on his approach. In Rawls's later work, he became especially concerned about the political stability of his theory of justice and sought to recast his theory in order to demonstrate how political stability could be achieved and maintained over time. This concern arose because Rawls recognized that for the foreseeable future, free individuals in a liberal democratic society will reasonably disagree about and endorse a range of different comprehensive philosophical, ethical, and religious doctrines. If people in a liberal democratic society will reasonably disagree about different comprehensive doctrines, then how can these individuals endorse and continue to support a common conception of justice? To solve this problem, Rawls recast his theory—Justice as Fairness—as a "political conception" of justice, which can be endorsed by an overlapping consensus of all reasonable comprehensive doctrines within a liberal democratic society. The core idea is that a "political conception" of justice is "freestanding" from any particular comprehensive doctrine and provides a framework in which members of a liberal society can offer public justifications to each other. However, while the political conception of justice is freestanding from any particular comprehensive doctrine, all reasonable comprehensive doctrines in a liberal democratic society will have their own internal reasons for endorsing this conception of justice. In this way, Rawls purports to show how Justice as Fairness can achieve political stability over time.

Following this Rawlsian strategy, Bai seeks to offer a "political conception" of Confucianism, which can be endorsed by an overlapping consensus of people who adhere to different comprehensive doctrines. However, Bai's approach differs in some significant ways from Rawls's approach. A key difference involves how the content of each theory of justice is determined. The content of Rawls's theory of justice is determined by using the Original Position, in which people are situated behind a veil of ignorance, unaware of certain aspects of their identity, and choose principles of justice in the absence of such knowledge. For Rawls, all reasonable members of a liberal democratic society endorse Justice as Fairness and agree that using the Original Position is the just

and fair way to determine the principles of justice (Rawls 1993, 22–28). Again, this is possible, according to Rawls, because each reasonable comprehensive doctrine in a liberal democratic society will have its own internal reasons for endorsing this, but they will all agree upon it through an overlapping consensus. It is important to notice that Rawls presents the idea of overlapping consensus *only* to show how his theory can achieve political stability over time, while it is the Original Position that is used to determine the content of his theory of justice.

In contrast, Bai determines much of the content of his theory of justice by relying on the works of the early Confucians. However, he also takes into consideration liberal democratic theory and what it may add. Bai describes this as showing how Confucianism can be “compatible” with liberal democracy. But despite this description, he seems to go further than exploring mere compatibility. It is not simply that political Confucianism is compatible with liberal democracy, but that elements of liberal democracy are incorporated into his Confucian theory. More specifically, Bai’s political conception of Confucianism offers what he calls a “hybrid regime,” which is a political system that incorporates meritocratic, democratic, and liberal elements. The Confucian hybrid regime seeks to limit nationalism and democracy—understood as “one person, one vote”—while embracing liberalism—understood as the rule of law and rights (244). In developing this model, Bai goes so far as to say, “. . . my critical proposal is in fact a support of (a revised version of) liberal democracy” (245). The suggestion is that Bai’s political conception of Confucianism is actually a version of liberal democracy, rather than merely compatible with liberal democracy.

But on what basis does Bai incorporate elements of liberal democracy into his Confucian theory, and how can he justify the claim that his theory is a “revised version” of liberal democracy? In developing the Confucian hybrid regime, Bai grounds his approach on the concept of overlapping consensus, which he explicitly attributes to Rawls. He explains his methodology as follows:

Rawls’s solution, simply put, is to take the whole theory of liberal democracy as a freestanding political conception, divorced from any known metaphysical ‘doctrine.’ This maneuver makes it possible

for different reasonable, liberal, or nonliberal doctrines to accept a common core, a political conception of liberal democracy that does not preclude the fundamental ideas of these doctrines. The *content* of liberal democracy is not predetermined by or derived from any a priori ideas but is an overlapping consensus worked out and endorsed by every reasonable and comprehensive doctrine.¹ (250, emphasis added)

While Rawls's theory may serve as the inspiration for Bai's approach, this passage reveals significant differences between them. First, as the passage indicates, Bai determines the content of his theory, and in particular its liberal aspects, by identifying an overlapping consensus of reasonable comprehensive doctrines (or more specifically, of Confucian and liberal democratic theory). Indeed, Bai states elsewhere that he will not follow Rawls in basing principles of justice on "equality, justice as fairness, or reciprocity" (254). We can understand this statement as indicating that Bai will not follow Rawls in using the Original Position to determine the principles of justice or content of his theory, and he instead employs the concept of overlapping consensus for this purpose. Second, Bai acknowledges that Rawls is concerned with how to achieve political stability, stating, "While sharing this concern with [Rawls], I am also concerned with helping people in a nonliberal democracy (as well as in a liberal democracy) to accept liberal democracy by showing that they can endorse liberal democracy and cherish their ideas that are different from and even in conflict with 'democratic' ideas" (252-53). In other words, Bai will use the idea of an overlapping consensus not merely to show how political stability might be achieved, but also to take into account both liberal and nonliberal comprehensive doctrines (such as Confucianism) to determine the content of his theory. As Bai states, in the passage quoted above, "The content of liberal democracy . . . is an overlapping consensus worked out and endorsed by *every* reasonable and comprehensive doctrine" (250, em-

¹ It is important to note that in this passage and others, Bai seems to attribute his own methodology to Rawls. Thus, it is not clear whether Bai recognizes that while he and Rawls may both appeal to the concept of overlapping consensus for certain purposes, they have very different methods for determining the content of their respective theories of justice.

phasis added). This is another significant deviation from Rawls, who directed his political conception of liberalism only to the members of liberal democratic societies. In other words, Rawls claims only that all reasonable comprehensive doctrines in a liberal democratic society will have reason to endorse Justice as Fairness and makes no claim that reasonable nonliberal comprehensive doctrines will endorse his theory. Furthermore, one of the things people are prevented from knowing in Rawls's Original Position is which comprehensive doctrine they endorse (Rawls 1993, 24; see also 25n27). In other words, for Rawls, comprehensive doctrines play no role in determining the content of the theory of justice, since a political conception of liberalism must be freestanding of any comprehensive doctrine. Rather, Rawls's political conception of liberalism derives its content from "fundamental ideas drawn from the public political culture of a democratic society" (Rawls 1993, 25n27). Third, Bai places a lot of emphasis on the idea of a "thin" conception of liberal democracy. The idea of "trying to 'thin down' liberal democracy in order for it to be inclusive" (251) is a direct result of basing the content of his theory on the common core of an overlapping consensus of every (liberal and nonliberal) reasonable comprehensive doctrine. Indeed, Bai acknowledges that his theory will be quite different from Rawls's conception of liberalism: "But my version of the common core of liberal democracy may explicitly be 'thinner' than Rawls's in certain respects, and only on the common core do I try to show that Confucianism is compatible with liberal democracy" (254). Given the very different methods for determining the content of their theories of justice, Bai's theory will necessarily be much thinner than Rawls's theory. It is not difficult to see that the common core of an overlapping consensus of every reasonable comprehensive doctrine will turn out to be much thinner than the content Rawls is able to derive from the Original Position.

While Bai's approach to political philosophy may be "inspired" by Rawls's approach, these contrasts should make it clear they are pursuing very different projects. Of course, one could challenge the premises of either theory, questioning whether the approach is plausible, will actually work, and so on. Here I have not tried to criticize or defend either approach, but instead to take each one at face value and highlight

the significant differences between them.

II. Bai on Human Rights

Let us now turn to the topic of human rights and explore how the methodological differences between Bai and Rawls bear on this aspect of their theories. Based on the overlapping consensus approach, Bai searches for a way to find a common core of agreement between the liberal concept of rights and Confucianism. While he acknowledges that the idea of rights is not to be found in the early Confucians (257), he argues that it is nevertheless possible to employ Confucian “strategies” for endorsing rights. Bai identifies three such strategies: “(1) replace rights talk with duties talk; (2) use the fallback apparatus; (3) refer rights to some higher good in Confucianism” (260). The first strategy relates to the fact that the early Confucians explicitly talk about duties, but not rights. This strategy involves placing an emphasis on obligations, rather than rights, so that rights are understood “not as the demandable right of the receiver but the demandable and (morally and legally) enforceable obligations and duties of the giver” (262). The second strategy relates to the fact that Confucians want people to willingly fulfill their obligations and believe that exerting too much coercion can be counter-productive to realizing that goal. However, despite this ideal, Bai claims that Confucians can still endorse moral and legal rights as a “fallback mechanism” that enforces certain duties when this is necessary. The third strategy involves justifying rights with reference to, and conditioning them on, higher goods recognized by Confucianism. These higher goods include ideals such as harmony, familial and communal care, and benevolent paternalism. This strategy will place conditions or limits on how certain rights are understood. For example, freedom of speech will not be interpreted as so robust that it allows neo-Nazis to march in a community of Holocaust survivors or permits pornography to be easily available (263), nor as permitting non-officeholders to engage in unlimited criticism of political officeholders (267). When the liberal concept of rights is combined with these Confucian “strategies,” we can identify an overlapping consensus

of liberal democratic theory and Confucianism, which yields a certain conception of rights. Bai argues that when rights are construed in this way, his political conception of Confucianism can endorse them and they can be incorporated into the Confucian hybrid regime.

Bai refers to the rights that result from this approach as “human rights.” He claims that his political conception of Confucianism can endorse human rights to freedom of speech, food, subsistence, education, healthcare, (267-68) and humane treatment or freedom from torture (277-78). He does not suggest that this list is exhaustive, so the political conception of Confucianism may be able to endorse additional human rights beyond those mentioned.

Bai certainly does not attempt to offer a comprehensive theory of human rights that would address many questions related to this topic. However, he does offer extended discussion of one specific dimension of human rights, namely, the use of human rights as standards for just war or foreign intervention. He completely rejects the idea of human rights providing standards for these actions, and instead advocates the concept of Confucian compassion. In order to understand Bai’s rejection of human rights serving this role, we must place this in the broader context of his approach to international relations.

Bai calls his general approach to international relations the “new tian xia model” (184), which is based on the virtue of Confucian compassion. Confucian compassion is a universal sentiment that essentially all people possess, at least to some extent. According to Bai, this virtue is the “social glue” (119) that can bond together a large society of strangers and provides the basis of his Confucian hybrid regime. However, it can also extend beyond the domestic state, to encompass other states, and even the entire world. Confucian compassion is hierarchical in nature and recognizes that one will have greater compassion for those who are near than for those who are more distant. The development of compassion or care for others begins in the family and must be extended step-by-step outwards. In this way, hierarchical care can be extended from the family, to the state, to the rest of the world (176).

The other key concept of the new tian xia model involves the distinction between civilized states and barbaric states. A civilized state will display civility through its culture, which does not have to be a

Confucian culture. There are at least two criteria that must be satisfied for a state to be civilized: (i) the legitimacy of the state lies in service to the people and the government is guided by the ideal of humaneness; (ii) this humaneness is reflected in its international conduct, such that it will never resolve conflicts with another civilized state through violence (186). A barbaric state, on the other hand, is “one that either tyrannizes its people or, out of incompetence or indifference, fails to offer basic services to its people, leaving them in great suffering; moreover, it threatens the well-being of other people or completely disregards its duty to other people, such as to protect a shared environment” (185). Tyrannical, failed, and ultranationalist states are mentioned as examples of barbaric states.

Combining the concept of Confucian compassion or hierarchical care with the distinction between civilized and barbaric states, the new tian xia model of international relations holds that “the people of one civilized state should ‘give preferential treatment to their own state over other civilized states,’ and people of all civilized states should ‘give preferential treatment to all civilized states over barbaric ones’” (Bai 185). It is important to note that preferential treatment does not mean indifference. For this reason, civilized states can intervene in the affairs of barbaric states, based on compassion for the people of the barbaric state.

We now come to Bai’s theory of just war and foreign intervention. He calls his position the “humane responsibility overrides sovereignty” view (227). This view builds on the new tian xia model of international relations, treating compassion or humaneness as the criterion for determining justified war and foreign intervention. Bai holds that it is best for civilized states to serve as a model of good conduct for barbaric states, as “the beacon on the hill,” and to reserve military intervention for only the most extreme cases (186). Nevertheless, the humane responsibility view treats a government’s right to sovereignty as dependent upon its humane treatment of the people. If the “state doesn’t practice humane governance but actively puts its people under unbearable misery, and if the ‘invaders’ are meant to save these people from their misery, then to defend the invaded state is fully unjust, while to welcome the invaders (liberators) is fully just” (227). In addition to

the requirement that there be a compassionate motivation to relieve the suffering of people, the humane responsibility view includes a few other important requirements: the suffering of the people must be to such a degree that they would welcome foreign intervention, the intervention must be endorsed by the international community,² and if the invading power is to remain in the invaded state, this must also be something that is welcomed by the people (227).

The “humane responsibility overrides sovereignty” view contrasts with what Bai calls the “human rights overrides sovereignty” view. He associates the latter with liberal theory. It treats respect for human rights as the basis of sovereignty, and violations of this duty as grounds for justified foreign intervention or war. Bai offers a number of criticisms of the human rights view, arguing that the humane responsibility view offers a better approach. First, he argues that foreign intervention on behalf of human rights can justify intervention in another society even when the people of that society do not recognize themselves as having such rights. The only thing necessary is that the intervening state recognizes the human rights in question. Among other problems, he argues, this can excuse and justify colonialism (230). In contrast, Bai claims that whether or not people are suffering is less controversial, and thus the standard used by the humane responsibility view provides a less problematic criterion for foreign intervention.

Bai’s contention that human rights can offer an excuse or justification for colonialism is not very convincing, because it overlooks the fact that colonialism itself would be a violation of certain human rights. A foreign state could not invade, set up illegitimate rule over a people, and claim fidelity to human rights. An additional problem with this criticism of the human rights view is Bai’s own discussion of cases that demonstrate the humane responsibility view can succumb to the same problem. More specifically, the humane responsibility view can also justify liberating people based on a value they do not recognize. He mentions the cases of the Chinese at the height of the Cultural Revolution and the current situation in North Korea, where

² Bai later qualifies the “international community” to include only civilized states, not all states (236).

people may be so oppressed and lacking in information, they believe their oppressed lives are fairly good. In these particular cases, Bai claims that the humane responsibility view may justify intervention, despite not all, or even a majority, of the people welcoming it (235). For both of the reasons mentioned, the human rights view does not appear to suffer from a problem that the humane responsibility view manages to avoid.

Bai offers a second criticism of the human rights view, arguing that it can justify foreign intervention and regime change even if the people of the target society are not ready for regime change. Furthermore, he argues, this could even have the problematic result of pushing the suffering people to side with their inhumane government and prolonging their suffering. Presumably, this problem is supposed to arise because violation of human rights is the only relevant consideration on the human rights view. Bai claims the humane responsibility view avoids this problem, because it requires that the people are suffering to such a degree, they would welcome the foreign intervention (231).

Once again, this criticism of the human rights view does not seem very convincing. First, on most developed theories of human rights, violations of human rights are a necessary, but not a sufficient, condition for justified foreign intervention.³ A state that seeks to intervene on behalf of human rights can exercise prudence and determine whether military intervention, or some other action, is the appropriate response to the violations. Bai implicitly recognizes this when he says that according to the human rights view, “a war of invasion in this situation *can* be just” (231, emphasis added), indicating that military intervention is not dictated by this view. Since the human rights view does not necessitate military intervention, it can take into account the probable reaction of the suffering population when determining the best overall course of action. If the people are not ready for a regime change, it may be very imprudent for a foreign power to carry out military intervention. This could involve a costly mistake, which requires the foreign state to stay and exercise some kind of governance, or risk having created a failed state. This type of situation may be illustrated by the United States’ second invasion of Iraq, a case that Bai

³ This point will be discussed in greater depth later in the paper.

mentions (228-29). Furthermore, while Bai is concerned with military intervention in these cases, it is important to note that military intervention is not the only available response to human rights violations. There can be a range of responses to such violations, which might take the form of diplomatic measures, economic sanctions, or military intervention, among others.

While these responses offer some reason to doubt that the human rights view fares worse than the humane responsibility view, there appear to be deeper issues at stake. First, it is important to notice that Bai's new *tian xia* model of international relations and the "humane responsibility overrides sovereignty" view of just war or foreign intervention are entirely derived from the early Confucians. There is no attempt to identify an overlapping consensus of Confucian and liberal democratic theory, which was the approach used to justify and incorporate human rights into the Confucian hybrid regime. We will return to this point later. Second, it is perhaps surprising to find Bai completely rejecting human rights as a standard for just war or foreign intervention. This surprise is due to Bai having followed Rawls in offering a "political conception" of justice, by drawing on certain ideas from Rawls's political liberalism to develop his political conception of Confucianism, but now rejecting Rawls's "political conception" of human rights. Rawls is the originator of not only political liberalism, but also of what has come to be known as a political conception of human rights. A key feature of political conceptions of human rights is their focus on the political role of human rights as norms that mediate the international relations between states.⁴ In order to better assess Bai's departure from Rawls on this particular point, we need to examine Rawls's conception of human rights and its role in his theory of international relations.

⁴ Political conceptions of human rights contrast with moral conceptions of human rights. While political conceptions focus on human rights as mediating the international relations between states, moral conceptions tend to focus on human rights as the moral rights that all people have simply in virtue of being human. Moral conceptions of human rights are more likely to embody what Bai calls a "moral metaphysics." For an example of a moral conception of human rights, see Griffin (2008). For other examples of a political conception of human rights, see Beitz (2009) and Raz (2010).

III. Rawls on Human Rights

While Bai refers to the rights endorsed by his Confucian hybrid regime as “human rights,” Rawls distinguishes between liberal rights and human rights, treating them as two distinct sets of rights. Liberal rights are the domestic rights of people within liberal societies, and are justified by the principles of justice chosen in the Original Position, as discussed above. In his later work, Rawls develops a theory of international justice, which governs the relations between “peoples” or societies. For Rawls, human rights are a “special class of urgent rights” that pertain to the international domain and play a key role in governing the relations between societies.

More specifically, Rawls claims that human rights specify “necessary conditions of any system of social cooperation,” where a system of social cooperation requires that members be given an adequate degree of respect and moral consideration (Rawls 1999, 68). Thus, these norms transcend liberalism and apply to all societies, both liberal and non-liberal. Rawls’s list of human rights includes the following:

the right to life (to the means of subsistence and security), to liberty (freedom from slavery, serfdom, and forced occupation, and to a sufficient measure of liberty of conscience to ensure freedom of religion and thought), to property (personal property), and to formal equality as expressed by the rules of natural justice (that is, similar cases be treated similarly) (Rawls 1999, 65).

In addition, he recognizes that minority groups have a human right to be secure from mass murder and genocide (79), and that women have human rights against oppression and abuse (75). Many commentators have noted that Rawls offers a rather short list of human rights, especially in comparison to the list found in the *Universal Declaration of Human Rights* (United Nations 1948) and other international human rights treatises. This short list of human rights is explained in part by the role they are given in Rawls’s theory of international relations.

Rawls calls his theory of international relations the “Law of Peoples.” This theory is based on the idea that peoples—or societies—are

equal, free, and independent, and that each decent society's freedom and independence should be respected by other decent societies. The Law of Peoples is comprised of eight principles. However, three of these principles, which pertain to war, intervention, and human rights, and are the most relevant for our purposes:

Peoples are to observe a duty of non-intervention (except to address grave violations of human rights).

Peoples have a right of self-defense, but no right to instigate war for reasons other than self-defense.

Peoples are to honor human rights. (Rawls 1999, 37)

As these principles reveal, just war is limited to self-defense and military intervention to addressing grave human rights violations. Thus, Rawls gives human rights a central and important role in international relations, since they serve as one of the primary norms that can justify military action.

The role of human rights is specified by three interconnected functions:

1. Their fulfillment is a necessary condition of the decency of society's political institutions and of its legal order.
2. Their fulfillment is sufficient to exclude justified and forceful intervention by other people, for example, by diplomatic and economic sanctions, or in extreme cases by military force.
3. They set a limit to pluralism among peoples. (Rawls 1999, 80)

For Rawls, human rights specify the limits of acceptable pluralism, or how much difference can be tolerated in the international community. If a society fulfills its human rights obligations, then it remains within these permissible limits and demonstrates itself to be a decent society. Furthermore, in demonstrating itself to be a decent society, it excludes itself from any justified intervention by other societies. On the other hand, if a society fails to fulfill its human rights obligations, and exceeds these permissible limits, then it demonstrates itself not to be a decent society. Furthermore, in demonstrating itself not to be a decent society,

it makes itself subject to justified intervention by other societies. Intervention in response to human rights violations may be diplomatic, economic, or military, with the caveat that military intervention is reserved for extreme cases.

Similar to Bai's categories of civilized and barbaric states, Rawls develops a typology of peoples or societies. A "people" is defined in terms of an institutional, cultural, and moral dimension. It involves a group of individuals that have a common government, shared sympathies, and a common conception of justice (Rawls 1999, 23-24). Rawls believes that peoples will seek to protect their political independence and territory, maintain their institutions and culture, and secure proper self-respect for themselves. There are five general types of peoples or societies: reasonable liberal peoples, decent peoples, outlaw states, burdened societies, and benevolent absolutisms (4). The first two types of peoples are decent, meaning they are peaceful and do not engage in unjust wars or aggression, have a common good conception of justice (one that aims at the good of its members), and honor the human rights of their people. Reasonable liberal peoples are peaceful and non-expansionist, governed by a liberal conception of justice, and honor the human rights of their people. Decent societies are peaceful and non-expansionist, governed by a nonliberal (common good) conception of justice, and honor the human rights of their people. The other three types of societies are not decent for one reason or another. Outlaw states violate the principles of international justice by being aggressive and expansionist and/or violate the human rights of their people. Burdened societies suffer from unfavorable conditions, which may be social or economic in nature, and are unable to maintain decent institutions or honor the human rights of their people. Burdened societies may include impoverished societies or failed states. Benevolent absolutisms may be peaceful and non-expansionist, and mostly honor human rights, but they do not allow their people a meaningful role in political decision-making, and thus fail to have decent political institutions.

Now let us turn to the issue of methodology. In order to justify the Law of Peoples and determine its content, Rawls introduces the idea of a second Original Position. While the first Original Position, discussed earlier, involves the members of a liberal society choosing principles of

justice to govern their own society, the second Original Position involves representatives of different societies deciding upon principles to govern their relations with each other. Rather than choosing among different principles of justice, as individuals did in the first Original Position, the representatives in the second Original Position are presented with eight principles, as mentioned above, and merely choose among “different formulations or interpretations” of those principles (Rawls 1999, 40). These eight principles are “familiar and largely traditional principles . . . take[n] from the history and usages of international law and practice” (Rawls 1999, 41).

It is significant that the Law of Peoples draws on the actual history and practice of international law, and merely allows the representatives to choose among different interpretations of these principles, rather than choosing among principles themselves. This is because in drawing on the actual history and practice of international law, Rawls exhibits a distinguishing feature of political conceptions of human rights. Political conceptions appeal to the practice of human rights, and especially to the political role of human rights in mediating international relations between societies or states. This appeal to practice has a formative influence in shaping political conceptions, because it directs their focus on the political role of human rights, rather than on, for example, the basic moral rights of persons.

The second Original Position involves a two-step process. The first step includes only the representatives of liberal societies, who decide which interpretations of the eight principles they will endorse. Rawls begins with this step because the “Law of Peoples . . . concerns what the foreign policy of a reasonably just liberal peoples should be” (Rawls 1999, 83). More specifically, the Law of Peoples is an extension of Rawls’s political liberalism to the international domain, and for this reason, it first and foremost determines the principles of foreign policy for liberal societies. However, Rawls claims that political liberalism includes a value of toleration, and this value necessitates a second step in the procedure.

Liberal societies recognize that, in the international sphere, equal peoples or societies will want to maintain their equality with each other and will also want due respect from other societies. Recognition

of this fact, in conjunction with the liberal value of toleration, makes it unreasonable for liberal societies to demand that all other societies also become liberal democracies (Rawls 1999, 59-62). To be clear, Rawls believes that only liberal societies are just. Nevertheless, some nonliberal societies, while not being fully just, are what Rawls deems decent, and should therefore be tolerated by liberal societies. These societies constitute the “decent peoples” mentioned above. Tolerating decent societies requires not only refraining from sanctioning them, but also treating them as “equal participating members in good standing of the Society of Peoples” (Rawls 1999, 59). Since decent peoples must be treated as equal members in the Society of Peoples, the second Original Position requires another step, which examines whether the principles of the Law of Peoples are acceptable from the perspective of decent peoples.

Rawls argues that decent peoples will accept the eight principles of the Law of Peoples. Since decent peoples are not aggressive and expansionist, and have a common good conception of justice, they will accept the (second) Original Position as a fair procedure for determining principles of international relations. Furthermore, since decent peoples have a common good conception of justice and protect the human rights and the good of their people, they will accept the principle of honoring human rights. In addition, since decent peoples are not aggressive or expansionist, they will accept principles that maintain a civil international order, such as the principle that permits war only in cases of self-defense and military intervention only to address grave human rights violations (Rawls 1999, 69).

It is important to notice that the liberal idea of toleration provides the basis for this second step. There is no appeal to comprehensive doctrines, since comprehensive doctrines are among the things representatives in the second Original Position are blocked from knowing. Rawls treats decent peoples as a type of nonliberal society with certain features, rather than as societies based on some particular comprehensive doctrine. Thus, Rawls’s use of a second Original Position to justify and determine the content of Law of Peoples continues to distinguish his methodological approach from the overlapping consensus approach used by Bai.

IV. Bai vs. Rawls on Human Rights

In order to assess Bai's and Rawls's competing views on human rights, it will be helpful to situate Bai's position within the Rawlsian framework of international relations. This requires us to more closely examine Rawls's concept of "decent peoples," because it can be shown that Bai's Confucian hybrid regime should be placed in that category. If this is the correct way to categorize the Confucian hybrid regime, then a representative of that regime will be included in the second step of Rawls's second Original Position, and we will need to check whether the Law of Peoples can be endorsed from the perspective of that representative.

Rawls suggests that there are two types of decent peoples. One type he calls "decent consultation hierarchies," while the other type is left unspecified, but supposes there may be some decent societies that do not fit the first model (1999, 63). A decent consultation hierarchy is "associationist," meaning it views individuals in society as members of groups, and these groups are represented by certain bodies in the political and legal system. A decent consultation hierarchy has a common good conception of justice that is typically grounded in a comprehensive religious or philosophical doctrine, and hence it does not have a political conception of justice like a liberal society. This common good conception of justice leads the society to respect and secure the human rights of its members, to have a legal system that is binding on all members of society, and to administer the legal system guided by the conception of justice. Furthermore, a decent consultation hierarchy is non-aggressive, pursues its aims through diplomatic means, and respects the independence of other societies (Rawls 1999, 64-67). As Bai acknowledges, the Confucian hybrid regime is not a liberal society according to Rawls's conception.⁵ However, we can also see the Confucian hybrid regime is not a decent consultation hierarchy, due to a couple of features it shares with liberal societies.

⁵ While Bai claims that the Confucian hybrid regime is actually a version of liberal democracy, he admits this claim requires one to embrace a broader conception of liberal democracy than the one offered by Rawls.

First, it lacks the “associationist” character of a decent consultation hierarchy. Like liberal societies, the Confucian hybrid regime does not treat its people as members of groups, but rather, as individuals. Second, like Rawls’s liberal society, the Confucian hybrid regime is based on a political conception of justice, whereas a decent consultation hierarchy is typically governed by a comprehensive conception of justice. So, the Confucian hybrid regime is neither a liberal society nor a decent consultation hierarchy. However, the Confucian hybrid regime does share the other features of a decent consultation hierarchy. It is a non-aggressive society, as shown by the new tian xia model of international relations and the limits this model places on war and foreign intervention. Furthermore, the Confucian hybrid regime involves a common good conception of justice, because “. . . the government is responsible for the material and moral well-being of the people. It is responsible for making it possible that average citizens have their basic material, social, moral, political, and educational needs met” (Bai 2020, 68). In order to ensure this aim, the Confucian hybrid regime seeks to elevate morally and intellectually superior people to political offices: “the right to participate in certain political activities is inseparable from one’s willingness to consider the common good and one’s competence at making sound decisions on this matter” (68). Finally, the Confucian hybrid regime honors human rights, as demonstrated by the Confucian “strategies” for justifying and incorporating human rights into this regime. So, while the Confucian hybrid regime is not a decent consultation hierarchy, these features place it among the other, unspecified, type of decent peoples.

Now that we have established Bai’s Confucian hybrid regime should be categorized as a decent society within the Law of Peoples framework, let us recall Rawls’s argument. The argument claims that due to the features of decent societies, they will endorse the second Original Position as a fair procedure for determining principles of international relations, and they will endorse the eight principles of the Law of Peoples as the appropriate principles for governing those relations. However, it is clear Bai will dispute this. Leaving the differing methodological approaches aside for now, Bai does not endorse the eight principles of the Law of Peoples. The Confucian hybrid regime endorses the new tian

xia model of international relations and the “humane responsibility overrides sovereignty” view of just war and foreign intervention. While the new tian xia model may share many similarities with the Law of Peoples, there is a clear difference between the “humane responsibility overrides sovereignty” view, which grounds just war and foreign intervention on Confucian compassion or humanness, and the human rights standard of foreign intervention offered by Rawls. Does this show that Rawls’s methodology has failed to justify his theory of international relations and the role he assigns to human rights?

Here it is worth noting that Rawls’s theory of human rights has been the subject of criticism, including criticism from liberal scholars. James Nickel (2006) criticizes Rawls for merely gesturing at the history and practice of international law and human rights, as shown by the eight principles Rawls offers for consideration in the second Original Position, while not considering thoroughly enough the contemporary practice of human rights. According to Nickel, this mistake leads Rawls to rely on grand dichotomies that oversimplify human rights and tie them too closely to being standards for foreign intervention. Instead, Nickel contends, contemporary human rights practice reveals that using human rights as standards for international intervention is only one among many roles that human rights can play. Standards for international intervention are not the central or primary role of human rights, and this role tends to apply only in the case of very severe human rights violations. Since international coercion and intervention can be “costly, dangerous, and often fail to work it is reasonable to restrict their use to the most severe human rights crises” (Nickel 2006, 271). Furthermore, Nickel points out that the human rights system places much emphasis on what he calls “jawboning,” which involves public “criticism or condemnation . . . that is not accompanied by significant threats” (271). Rather than making coercion and intervention the central response to human rights violations, we find that “many human rights treaties deal with human rights violators through gentler means, such as consciousness-raising, persuasion, norm-promotion, criticism, shaming, defining conditions for full acceptance, mediation, and negotiation” (273). Given these features of human rights practice, Nickel suggests a better description of the main role of human rights

it that “they *encourage and pressure* governments to treat their citizens humanely. . . .” (271, emphasis in the original).

We should also consider the important work of Charles Beitz (2009), who offers perhaps the most developed version of a political conception of human rights. Like Rawls, Beitz appeals to the history and practice of international human rights to develop his theory. However, as Nickel suggests, Beitz goes further than Rawls, explicitly grounding his theory in a full account of contemporary human rights practice (Beitz 2009, chap. 2). Like Nickel, Beitz recognizes a range of roles that human rights can play. This leads Beitz to propose a “two-level model” of human rights. At the first level, human rights apply to the domestic political institutions of states, “including their constitutions, laws, and public policies” (109). At this level, human rights create responsibilities for governments to respect, protect, and fulfill the rights of people within their state. At the second level, human rights are “matters of international concern” (109). The international concern arises when a government fails to meet its first-level human rights obligations, which provides *pro tanto* reasons for capable outside agents, including other states and the international community, to act. These actions can include holding states accountable for meeting their human rights obligations, assisting states that lack the ability to meet their human rights obligations, and intervening in a state to protect human rights (109). Beitz contends that the international role of human rights is “. . . perhaps the most distinctive feature of contemporary human rights practice” (115), and that “. . . the interference-justifying role is central to understanding their discursive function” (116). Thus, Beitz recognizes not only a range of roles that human rights can play, but also a range of actions that may be justified when a government fails to meet its first-level human rights obligations. Foreign intervention is one among the possible responses to such a failure. However, Beitz also treats these responses in a more nuanced manner, by construing human rights failures as providing *pro tanto* reasons for outside agents to act, meaning violations do not provide conclusory reasons for action (including intervention), because these reasons must be weighed against competing reasons.

If we take into account Nickel’s criticisms of Rawls and Beitz’s more developed political conception of human rights, we may identify ways in

which Rawls's theory of human rights should be revised. Human rights can serve a range of different roles, and they need not be tied so closely to the particular role of standards for foreign intervention. Instead, human rights can serve as norms that enable societies to "encourage and criticize" other societies, and this need not involve threats or intervention. In other cases, a failure to meet human rights obligations may be due to a lack of resources, in which case the appropriate response could be foreign assistance, not intervention. By recognizing a wider range of roles for human rights and tying them less closely to standards for foreign intervention, Rawls would be able to recognize a longer list of rights, such as rights to education and healthcare, which are included in Bai's list of human rights. This is because when a state fails to provide healthcare or education for its people, it is typically not cause for foreign intervention, although it might be cause for foreign criticism, encouragement, aid, or other forms of assistance.

While these revisions would allow Rawls's theory to more accurately and fully reflect the practice of human rights, it is important to note that neither Nickel nor Beitz rejects the role of human rights as providing standards for just war or foreign intervention. Nickel suggests foreign intervention should only be an appropriate response to the most severe human rights violations, while Beitz recognizes foreign intervention as among the possible responses to human rights violations, tempered by a consideration of the competing reasons for other forms of action (or inaction). So even when taking these criticisms and possible revisions into account, Rawls's theory of human rights would still recognize a role for human rights to provide standards for just war or foreign intervention.

Having established that a revised version of Rawls's theory will still recognize a role for human rights to provide standards for just war or foreign intervention, we can return the issue of how Bai, or the representative of the Confucian hybrid regime, should respond to the Rawlsian framework. Suppose Bai continues to object to this role for human rights because it embodies the "human rights overrides sovereignty" view, whereas the Confucian hybrid regime endorses the "humane responsibility overrides sovereignty" view. For this reason, Bai argues that the Law of Peoples is not acceptable to all decent

societies, based on Rawls's own methodology of the second Original Position. We now find the discussion coming back around to the issue of methodology.

It does not appear Bai can justify this rejection of the Law of Peoples and the role that Rawls assigns to human rights. The problem is not due to Rawls's methodology, but rather, because Bai violates his own methodology. Recall that Bai justifies and incorporates human rights in the Confucian hybrid regime by appealing to an overlapping consensus of Confucianism and liberal democratic theory. This overlapping consensus involves combining the liberal concept of rights with the three Confucian strategies for endorsing rights. However, as briefly discussed above, when Bai develops the new *tian xia* model of international relations, with its "humane responsibility overrides sovereignty" view, it is derived purely from the works of the early Confucians. There is no attempt to identify an overlapping consensus of Confucianism and liberal democratic theory. Instead, Bai completely rejects the idea of human rights as providing standards for just war or foreign intervention, or as he calls it, the "human rights overrides sovereignty" view, in favor of the Confucian-inspired "humane responsibility overrides sovereignty" view.

It is important to notice that this is not merely an obscure methodological point but goes to the very heart of Bai's theory. Bai acknowledges that rights are not to be found in the work of the early Confucians. It is only by employing the overlapping consensus approach, which identifies an overlapping consensus of liberal democratic theory—and specifically, its concept of rights—and Confucianism, that Bai justifies the inclusion of human rights in the Confucian hybrid regime. If Bai is to consistently apply this methodology, he must also look for an overlapping consensus in the area of international relations. This involves finding a common core of agreement between the Confucian concept of compassion or humaneness and human rights, or in other words, between the "humane responsibility overrides sovereignty" view and the "human rights overrides sovereignty" view. Perhaps human rights could serve as *précising* norms for Confucian compassion or humaneness. That is, while Bai could maintain the general approach of using humane responsibility as the standard for just war or foreign

intervention, the violation of human rights would determine precisely when the mistreatment of people has reached such a level that compassion or humaneness demands foreign intervention. Adopting this approach, Bai would not only adhere to his own overlapping consensus methodology, but it would also allow his political conception of Confucianism to incorporate a political conception of human rights, one which recognizes the contemporary practice of human rights and the political role of human rights in mediating international relations between societies or states.

However, while this proposal might seem to resolve the issue, the problem actually goes deeper. Recall that Bai's methodology requires him to identify an overlapping consensus "endorsed by every reasonable and comprehensive doctrine" (Bai 2020, 250). Even if Bai adopts the suggestion offered above, he has merely identified an overlapping consensus of Confucianism and liberal democratic theory. He tells us that his attention is focused on an overlapping consensus of Confucianism and liberal democratic theory because liberal democracy has an "end of history" status (1; 97). But the problem is that there is no reason to believe reasonable comprehensive doctrines are limited to Confucianism and liberal democratic theory. For example, while still confining our focus to the issue of human rights, consider the *African Charter on Human and Peoples' Rights* (Organization of African Unity, 1981). When this regional human rights treaty was developed by the Organization of African Unity (later replaced by the African Union), they specifically included not just individual human rights, but "peoples" rights, which are group rights. The inclusion of "peoples" or group rights distinguishes the African Charter from other regional human rights treaties, and this was felt to be important because it reflects the communal aspects of African society and thought. Thus, the inclusion of "peoples" rights is reflective of comprehensive doctrines found in African societies. Presumably, these are reasonable comprehensive doctrines.⁶ If Bai's methodology requires that he identify an overlapping

⁶ Here we should recall that Rawls recognizes decent consultation hierarchies as being reasonable societies. In the case of decent consultation hierarchies, every individual is viewed as a member of a group, and groups receive representation in the political institutions. The *African Charter on Human and Peoples' Rights* is certainly closer to

consensus of every reasonable comprehensive doctrine, both liberal and non-liberal, then his consideration of Confucianism and liberal democratic theory has only partially completed the work that must be done. Identifying an overlapping consensus with other reasonable comprehensive doctrines, such as those found in African societies, will also be necessary. Taking these comprehensive doctrines into account, including the idea of communal or group rights, will almost certainly require further modification of Bai's theory of human rights.

V. Conclusion

In conclusion, we can see that Bai has not fully adhered to his own methodology when developing a theory of human rights. This methodology requires that he more fully take into account the liberal theory of human rights. More specifically, this must include recognition of the role that human rights play in the international domain and meditating the relations between states, including their role as standards for just war or foreign intervention. This step would perhaps bridge differences between Bai's and Rawls's theories of international relations and human rights, so that the representative of the Confucian hybrid regime could endorse something approximating the Law of Peoples. However, the demandingness of Bai's methodology is revealed when we recognize that there are other reasonable comprehensive doctrines, besides just Confucianism and liberal democratic theory. Since Bai bases his approach to political philosophy on an overlapping consensus of every reasonable comprehensive doctrine, these other reasonable comprehensive doctrines must also be taken into account when determining the consensus. This will almost certainly require further modification of his theory.

liberalism than a decent consultation hierarchy, because it simply recognizes communal or group rights, in addition to individual rights. Therefore, Rawls would almost certainly recognize the comprehensive doctrines reflected in the *African Charter* as reasonable.

REFERENCES

- Bai, Tongdong. 2020. *Against Political Equality: The Confucian Case*. Princeton, NJ: Princeton University Press.
- Beitz, Charles R. 2009. *The Idea of Human Rights*. Oxford, UK: Oxford University Press.
- Griffin, James. 2008. *On Human Rights*. Oxford, UK: Oxford University Press.
- Nickel, James W. 2006. "Are Human Rights Mainly Implemented by Intervention?" In *Rawls's Law of Peoples: A Realistic Utopia?*, edited by Rex Martin and David A. Reidy, 263-77. Hoboken, NJ: Wiley-Blackwell.
- Organization of African Unity. 1981. *African Charter on Human and Peoples' Rights*. Accessed January 1, 2022. <https://www.achpr.org/legalinstruments/detail?id=49>
- Rawls, John. 1993. *Political Liberalism*. New York City: Columbia University Press.
- _____. 1999. *The Law of Peoples*. Cambridge, MA: Harvard University Press.
- Raz, Joseph. 2010. "Human Rights Without Foundations." In *The Philosophy of International Law*, edited by Samantha Bisson and John Tasioulas, 321-38. Oxford, UK: Oxford University Press.
- United Nations. 1948. *Universal Declaration of Human Rights*. Accessed January 1, 2022. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>.

■ Submitted: 20 Nov. 2021

Accepted: 16 Feb. 2022