A Dialectic of Justice: Critiquing Social Conventions and Motivating Practice

Samuel Webb, Tufts University, Class of 2010

Socrates’ infamous question, “What is Justice?” in Plato’s Republic, spawned a debate for philosophers (and societies), and the answers continue to be important and contested today. The diversity of approaches and persistence of disagreement has inclined some contemporary philosophers to adopt a position of moral relativism. Relativism is the thesis that there are many equally valid frameworks for making value judgments and that these judgments are right or wrong only relative to the operative framework. One advocate of this position, Gilbert Harman, argues that our various moral frameworks are ‘completely conventional’, the result of a social contract whose terms arise from the self-interested bargaining of its participants (Harman 6). In contrast to this approach, John Rawls offers a direct and non-relative response to Socrates’ problematic question: the theory of ‘justice as fairness’. Rawls claims that, although particular conceptions of justice may vary from society to society, the concept, i.e., the criteria governing what counts as a conception of justice at all, is non-conventional. As such, it is not subject to the kind of bargaining described by Harman.

This paper aims to show that Rawls’ theory raises a significant challenge to Harman’s account of morality. By making conventional social agreements the basis of morality, Harman effectively removes the possibility of evaluating those conventions. But such conventions, because they arise from bargaining among parties of unequal power, fail to meet Rawls’ ‘uncompromising’ criteria of justice as fairness and can thus be critiqued precisely on the grounds of justice. To that end, Rawls holds that, through hypothetical deliberation behind a ‘veil of ignorance’, one can arrive at principles of justice which are in conflict with the existing conventions, even if those principles are not in one’s own interest. One reason for Harman to reject such a conception of social justice, I will suggest, is that he thinks that it cannot motivate actual social practices. In this sense, I will argue, Harman is fundamentally making a descriptive claim based on his doubt that Rawls’ normative claim (that unjust social agreements should be rectified in light of considerations of fairness) is practical.
First, I will explain in greater detail the two positions and how they come into conflict; next, I will address Harman’s concern that existing social agreements might be unjust; and finally, I will consider how the two positions may be reconciled to enhance our ability to critique existing conventions. The guiding aim of my discussion is to show that any adequate conception of justice needs this critical capacity, but must also address the challenge of motivating people to practice its principles in light of actual power relationships in society.

Harman’s account begins with the suggestion that most people’s values “reflect conventions that are maintained by continual tacit bargaining and adjustment” (Harman 22). This claim seems innocuous enough until it is coupled with the stipulation that such values (including justice) arise entirely from the continual process of compromise between the different self-interested individuals and power groups in society. Customs arise because they are useful for social cooperation and these include everything from simple formalities to the deepest moral convictions that govern people’s attitudes and actions. They must be advantageous enough for everyone to accept initially, but once accepted they solidify and reflect the relative bargaining power of those in the agreement. Harman’s paradigmatic case is the difference in moral weight given to the prohibition against harm over the obligation to help. This is meant to illustrate how the existing moral code reflects the interests of the more powerful members in society who are less likely to be in need of help, but who benefit as much or more than others from the prohibition against harm.

Rawls might be willing to grant Harman that some values and moral judgments are indeed derived from actual bargaining in society. Yet, justice itself cannot be such a value. For a concept of justice precedes and frames the agreement (and disagreement) on which society is based; it is the ‘first virtue of social institutions’ (Rawls 3). Every society, even one in which the particular conception of justice is hotly contested, insists Rawls, “understands the need for [and is] prepared to affirm a characteristic set of principles for assigning basic rights and duties and for determining what they think should be the proper distribution of the benefits and burdens of social cooperation” (Rawls 3). The role these principles play is the non-relative concept of justice.

Of course, different societies and members of societies will disagree about which particular principles and distributions should be chosen. On this point, Rawls does not dispute the relativist. The key difference for Rawls, however, is that the “rights secured by justice are not subject to political bargaining or the calculus of social interests” (Rawls 3). Instead, justice demands that all other values, institutions and agreements be measured against those that “free and rational persons would choose in an initial position of equality as defining the fundamental terms of their association” (Rawls 10). It is in this sense that justice can be understood as fairness; it is conceived without the advantage of any particular persons in mind. In short, it is the limiting principle which permits the evaluation of further agreements and institutions once the contract has been established.
For Rawls, the hypothetical nature of this contract is what enables it to capture the ideals of justice. Certainly actual societies were not founded on the voluntary association of equal persons. Differences in power, interests, capacities, class, conceptions of the good, living conditions etc. have contributed to the form of existing conventions, rendering them, to some extent, unfair. For this reason, justice as fairness must abstract from these differences:

The principles of justice are chosen from behind a veil of ignorance. This ensures that no one is advantaged or disadvantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances. Since all are similarly situated and no one is able to design principles to favor his particular condition, the principles of justice are the result of a fair agreement or bargain. (Rawls 11)

It is important to remark that while Rawls uses the same terminology as Harman, that of a ‘bargain’, what Rawls means does not fit Harman’s picture. The veil of ignorance ensures fairness but it also eliminates the possibility of ‘bargaining’ as such, since no one is advancing particular interests. What Harman describes is the real process of coming to a compromise over conflicting interests. Rawls’ concept of a ‘fair bargain’ is better thought of as a form of collective deliberation. The agreement reached must be one to which all members would voluntarily assent regardless of their bargaining power, a possibility which Harman rejects. For all these reasons, it would seem that the actual agreements described by Harman are unjust and consequently ill-suited to be the alleged source of the moral obligations and convictions of which justice is a part.

Harman addresses this objection directly. He acknowledges that, “It might be thought that there is something unjust about agreements whose outcomes are affected by the differences in power relationships among the affected parties” (Harman 28). This would be deeply problematic because according to Harman’s brand of relativism, such judgments about justice must be based on values themselves derived from moral conventions accepted by those who make them. This might point to a “tension or instability between certain moral views and an explanation in terms of differences in bargaining power” (Harman 28). Giving up the conventional explanation would jeopardize Harman’s whole project and leave the source of many other moral judgments uncertain. But is it possible for a conventional moral judgment to be at odds with the power arrangement that yields it? How could our conception of justice have come from an agreement that identifies itself as unjust?

These questions should make us skeptical that Harman’s conventionalism can adequately account for justice, but we should not hastily abandon his insights about power relationships in society. The rich history of thought on justice can help us here. Blaise Pascal, speaking from a radically different political and ethical context, offers an enlightening perspective:

No doubt the equality of goods is just, but... Being unable to force one to obey justice we have made it just to obey force. Unable to fortify justice we have justified force so as to unite the just and the mighty so that there should be peace, which is the sovereign good. (Pascal f. 76)
It is possible to identify or even agree upon a conception of justice without it having any normative or motivational weight for us. Let us call this Pascal’s problem. It consists in recognizing a principle of justice (in this case, quite an egalitarian, distributive one) which nonetheless remains impotent or impractical. Unlike formal truths, which we need only grasp rationally in order to accept, we must be inclined towards justice in order to follow it. In other words, we might recognize the claims of justice without accepting and being motivated by them.

Force, on the other hand, compels obedience by itself. Exactly because principles of justice do not serve the interests of the powerful (whether it is under the Ancien Régime or liberal capitalism), and because they are public, they are susceptible to being masked and melded with force in order to achieve ‘peace’ or the stability of the contract. Like Rawls, Pascal recognizes a conception of justice that transcends society’s power structure, but he contends, with Harman, that such a conception cannot meaningfully operate while power is a determining factor in the social contract. In fact, Pascal believes true justice lies beyond humanity’s reach, but we will use his problem rather more optimistically as a tool to evaluate theories of justice in terms of their potential to motivate adherence to their principles.

The union of justice and force is, in a manner of speaking, one of the ways in which Harman suggests resolving the tension in his theory. Perhaps an agreement based on force is just after all; contracts between large banks or corporations and individuals are generally considered fair, despite the disparities in bargaining power, he comments. This seems like a rather transparent example of what Pascal’s problem laments—allowing force to be incorporated into the idea of justice such that it essentially replaces it. The other option is to:

[…]
Concede that the unfairness of the agreement provides some reason to try to modify the results of the agreement so that they come to be more in accord with what might have emerged if the bargaining position of the affected parties had been the same. (Harman 28-29)

If he were to pursue that option, to acknowledge that principles of justice could be active in shaping what results from the bargaining, Harman would seem to concede to Rawls’ position. However, unlike Rawls, he still couches his analysis in terms of the likely attractiveness to individuals actually bargaining. Anyone advocating the above perspective on existing agreements would disturb the consensus, allowing for the potential institution of new moral conventions, but the problem of disparate bargaining power would remain. Ultimately, Harman’s attachment to moral conventionalism prevents him from adequately addressing the possibility that people in positions of bargaining power might develop a conception of justice not calculated to serve their own interests. This seems an impoverished view of human motivation, one which totally overlooks genuine concern for others. Faced with precisely the kinds of justice proposed by Rawls or in Pascal’s problem, Harman appears to be, in some sense, begging the question.
Yet, we should be careful not to caricature Harman’s position. He is not simply declaring, as Thrasymachus does in *The Republic*, that “justice is the advantage of the stronger” (Plato 338c). Rather, he is observing that the stronger, i.e., the socially powerful, can and do *use* their advantage in bargaining for social conception of justice. One objection he might legitimately make to Rawls is that the ‘fair agreement’ and the ‘veil of ignorance’ are just wishful hypotheticals. They are not relevant to the actual determination of consensus. Moral conventionalism accounts for actual agreements as they seem to occur, not how they would ideally occur. Real agreements only appear ‘unjust’ from the perspective of the disadvantaged bargainer when he or she rejects the contract in an effort to reconstitute it. It is Rawls who is failing to acknowledge Pascal’s problem, Harman might argue. Such a claim cuts to the fundamental difference between their two projects: Harman seeks primarily to describe how moral judgments are commonly formed while Rawls would like to offer a model as to how they might be formed better.

Rawls is himself aware of the tension, however:

> It is natural to ask why, if this agreement is never actually entered into, we should take any interest in these principles moral or otherwise. The answer is that the conditions embodied in the description of the original position are the ones we do in fact accept. Or if we do not, then perhaps we can be persuaded to do so by philosophical reflection. (Rawls 19)

In order for a hypothetical agreement to justify its principles, it must in some way represent conditions that we actually see as necessary for ensuring an agreement is fair. Rawls is confident that his description of the original position as a situation of equality behind the veil of ignorance (“where the outcome [of agreement] is not conditioned by arbitrary contingencies or the relative balance of social forces” (Rawls 104)) fulfills our criteria for fairness. If it does not, Rawls suggests that in considering the problem posed by the initial situation, we can cash out or reveal *existing but unacknowledged* judgments about what makes an agreement just. We may even be led to revise those judgments by reflecting on the conditions he suggests. This point is crucial because Rawls needs to persuade us that these restraints on the choice of principles are not arbitrary and artificial but reasonable and natural.

Furthermore, if we are to be persuaded to accept these conditions of fairness, we must be able to see how they could come from us and how we might share them. The young Marx strikes a similar chord:

> We develop new principles for the world out of the world’s own principles. We do not say to the world: “Cease your struggles, they are foolish; we will give you the true slogan of struggle.” We merely show the world what
it is really fighting for, and consciousness is something that it has to acquire, even if it does not want to.

(Marx Letter)

Although Marx would have been unsatisfied with Rawls’ non-revolutionary political philosophy, the two share the method of what we might call *revelatory persuasion*. Principles of Rawlsian justice are not abstract absolutes imposed from on high; they are expositions of the existent *possibilities* for social cooperation. By deriving our principles from a hypothetically fair agreement we expose actual bargains as unfair and can critique any conventions arising from them. Our confidence in conventional moral judgments should be undermined, especially if justice as fairness provides a possible alternative we are more inclined to accept. This allows for reflective deliberation on principles to replace bargaining. Rawls said it in the following way: “We need a conception that enables us to envision our objective from afar” (Rawls 19). Rawls’ hope is that if his theory is not merely descriptive but revelatory, we will find his picture of the original position compelling, and so will be convinced by collective reflection to freely impose its corresponding conception of justice on ourselves.

Harman, I think, has much less faith in the potential for philosophical reflection to enlighten us on the principles of justice as something over and above self-interested compromise. He is motivated in this view in part, as I have said, by his commitment to explaining morality exclusively through conventions. This concern alone, though, seems shaken in the face of the Rawlsian challenge. His more reasonable objection stems from implicitly recognizing Pascal’s insight. In order for justice to be a functional concept, the powerful must be inclined to act in accordance with it even when it does not serve their interests. Harman’s account leaves no room for a principle of justice to impose itself by its own sheer force, even on reflection. If an individual or group is discontented with a given social arrangement or moral convention, they may invoke fairness to reject the existing contract in an attempt to reform it to their advantage, disrupting ‘peace which is the sovereign good’, but the process will remain one of bargaining against force. What is central for Harman is that such action can only be reliably motivated by the interests of the parties involved, not by accepting a principle that transcends their interests.

For this reason, Harman’s account of justice alone is neither complete nor satisfying. There is another way to interpret Rawls’ theory in terms of Pascal’s problem that preserves its revelatory critical power while addressing the effect of force on motivating its principles. Although Rawls and Harman disagree on how values can be constituted in society, neither looks to *a priori* values that exist over and above society. They are both social contract theorists and as such share an appreciation for the fact that agreement is essential to constitute values of any kind. This is crucial because it suggests that their accounts need not be mutually exclusive. When Rawls asserts that the concept of justice is foundational for society (not the result of bargaining already within society) he also acknowledges that it has other key features to consider, such as coordination, efficiency and stability. This is where the most room opens up in his approach to find common ground with Harman.
Rawls assumes that a conception of justice selected from behind the veil of ignorance will be stable. Notably, this means that ‘the principles of moral psychology’ will not “fail to engender in human beings the requisite desire to act on it” (Rawls 398). Yet, Harman seems convinced that principles of the type reached through the veil of ignorance device do not have the same motivating force as conventionally established moral judgments. How can we reconcile this? Recall Pascal’s problem: any conception of justice that transcends the distribution of force in society will necessarily be unstable not because it conflicts with abstract principles of moral psychology but because of the particular, concrete power relations.

It is possible that ‘justice as fairness’ arose conventionally, via bargaining, to be a conception that transcends and rejects its origin in self-interest. Rawls opens up the roots of justice for explicit philosophical critique beyond tacit bargaining. Yet, whatever conception is established this way (i.e., by critical philosophical reflection and hypothetical agreement) still faces two problems—not only that of breaking an existing framework reinforced by real, self-interested bargaining, but also of constantly rechecking itself against the persistent eroding effect bargaining has on conceptions of justice. Rawls says that stability is needed for a ‘viable human community’ (Rawls 3), but he mistakenly assumes this means deciding “once and for all what is to count […] as just and unjust” (Rawls 11). If we take Pascal and Harman seriously, this is not possible. A theory of justice cannot be static once established in a given society. It must be continually self-critical to avoid tacit modification from the force of bargaining and changing circumstances. We need to be reminded, lest we forget, what we are fighting for. Thus, the challenge that lies before us becomes not what justice is or how we came to that conception, but how we can overcome Pascal’s problem of its real, ongoing application.

---

i See the two-part definition (Harman 4-5).

ii Harman’s view appears vulnerable to a Marxist critique here. If morality is the result of self-interested bargaining reflecting the unequal power structure of society, then it seems hard not to concede to the Marxist who argues that morality is merely an ideological cover for the domination of one class by another, in which case its legitimacy is highly dubious, especially to the dominated class. This topic cannot be fully addressed here.

iii Rawls’ hypothetical “original position” takes the place of the “state of nature” of traditional social contract theory.

iv In searching for the appropriate description of the original position, Rawls proposes the concept of “(wide) reflective equilibrium” (p. 18, 40-5) to describe a state in which, by mutual consideration and adjustment, our moral judgments and principles coincide. This idea has been much analyzed: cf. Norman Daniels “Wide Reflective Equilibrium and Theory Acceptance in Ethics” The Journal of Philosophy, 76.5 (1979): 256-282 and Harman’s own 2003 paper “Three Trends in Moral and Political Philosophy”. The Journal of Value Inquiry. 37.3: 415-425. Notably, while it involves reciprocal “adjustment” like Harman’s account, reflective equilibrium definitively lacks the aspect of bargaining Harman identifies as constitutive of our values.

v “Sans doute, l’égalité des biens est juste, mais… ne pouvant faire qu’il soit force d’obéir à la justice, on a fait qu’il soit juste d’obéir à la force; ne pouvant fortifier la justice, on a justifié la force, afin que le juste et le fort fussent ensemble, et que la paix fût, qui est le souverain bien.” This is my translation.
Pascal’s thought on justice is complex and controversial, but we might characterize it by the idea that due to humanity’s fallen condition, we can discern only traces of divine justice for whose absence we institute conventional, human justice by force and deception. Cf. the Pensées, fragment 56, (294 Br., 60 Laf.).

Indeed, he would seem to be closer to Rawls than to relativism. I am indebted to Professor Erin Kelly for pointing this out.

Despite his attempt to characterize morality as it actually is, if our moral lives and personal relationships were as predicated on self-interest as Harman implies, would we recognize them? What would we make of the sacrifice of parents for their children or romantic love? To interpret certain important areas of human experience and commitment on the basis of self-interested is to distort them.

See Rawls 42-43.

Rawls argues that “a society satisfying the principles of justice as fairness comes as close as a society can to being a voluntary scheme […] In this sense, its members are autonomous and the obligation they recognize self-imposed” (Rawls 12). This fulfills the Kantian ideal of morality as laws we give ourselves autonomously and suggests a source of motivation.

Although Rawls is in some sense a Kantian, he maintains that justice is not a ‘necessary truth’ nor can it be “deduced from self-evident premises or conditions on principles,” (Rawls 19) (as a pure Kantian might) but he nonetheless believes conceptions of justice engender a corresponding sense of justice, and, hence, motivational weight derived not from our particular desires or interests but from reaching wide reflective equilibrium. Principles for a relativist like Harman, as for Humeans, possess no inherent motivational force.
Bibliography


Norman, Daniels. “Wide Reflective Equilibrium and Theory Acceptance in Ethics.” The Journal of

Philosophy 76.5 (1979: 256-282.


Plato. Republic.