ON RAWLS' JUSTIFICATION PROCEDURE

Richard A. O'Neil

March 10, 1976
Abstract:

The paper is a defense of the moral methodology of John Rawls against criticisms by R.M. Hare and Peter Singer. Rawls is accused of intuitionism and subjectivism by Hare and of subjectivism and relativism by Singer. I argue that Rawls does not rely on intuitions as such, but on judgments on which there is a consensus. This does not commit Rawls to subjectivism for what is required for objectivity in ethics as in science is simply a rational justification procedure for principles, which Rawls provides. Moreover, an appeal to a moral consensus at some point is inescapable. Finally, concerning the charge of relativism, I point out that Rawls includes in his justification procedure only those judgments on which there is a consensus among competent judges. Though there is the possibility that conflicting sets of judgments may be equally valid for different societies in the unlikely case that there is nothing invariant in the judgments of competent judges, this is a relativism we can accept.
On Rawls' Justification Procedure

In discussions of his moral methodology, John Rawls has been described as an intuitionist and subjectivist by R. M. Hare\(^1\) and as a subjectivist and relativist by Peter Singer.\(^2\) Since these descriptions are applied in criticism of Rawls it is important to determine whether they are appropriate. This is not an easy task, since philosophers have used each of the terms to refer to a variety of types of philosophical theory. Rawls might be an intuitionist, a subjectivist, or a relativist in one sense of these terms but not in other senses. In this essay I will argue that while there is a sense in which each of these terms may be applied to Rawls, the views on the basis of which they are applicable are justified. Thus I will support Rawls' methodology against the criticisms of Hare and Singer.

I

Rawls uses the term "intuitionist" for one who holds there is an irreducible plurality of first principles of justice and no method for weighing them against one another except that of appeal to "what seems to us most nearly right."\(^3\) I shall disregard this rather idiosyncratic use of the term since Rawls argues against this view and since this is not the sense in which Hare claims Rawls is an intuitionist. Hare says that Rawls "certainly is one in the usual sense" and cites Rawls' own admission that his theory appeals to intuition.\(^4\) In addition, Hare indicates several instances where he believes Rawls relies on intuitions to support his argument. Unfortunately, however, Hare fails to specify which is "the usual sense" of


\(^3\)\textit{A Theory of Justice} (Cambridge, Mass., 1971), p. 34. Henceforth abbreviated in the text to \textit{ATOJ}.

\(^4\)Hare, "Rawls' Theory," p. 146.
"intuitionist," and so before we can evaluate Hare's claim an investigation into the way in which the term has been used is necessary.

As Rawls notes (ATOJ, p. 34), intuitionism in its traditional (and, I believe, usual) sense has included certain epistemological theses concerning the nature of moral principles and the manner in which we acquire knowledge of them. Intuitionists of this sort believe (a) certain moral principles are self-evident or necessary, and (b) these principles are known to be true through direct rational insight: provided we understand them we just "see" or rationally apprehend them to be true. Some intuitionists claim there is just one such principle (Sidgwick) while others believe there are several (Ross and Price).

Traditional intuitionism encounters a number of difficulties. For example, it must make sense of the notion of a self-evident moral truth and it must provide a test for the presence of genuine intuitions. Thus Hare would have grounds for criticizing Rawls if Rawls were an intuitionist of this type. However, Rawls holds neither of the epistemological theses mentioned above: (a) He clearly rejects the notion of self-evident or necessary moral truth. "There is no point at which an appeal is made to self-evidence in the traditional sense either of general conceptions or particular convictions" (ATOJ, p. 21). (Cf. Singer, who suggests that the proper way to proceed in ethics is to "search for undeniable fundamental axioms."5) (b) Rawls cannot hold the view that the manner in which we know the truth of some principle is through intuition for Rawls stresses the fallibility of intuitive judgments. "...Even the judgments we take provisionally as fixed points are liable to revision" (ATOJ, p. 20). If intuitions can be mistaken, I cannot claim that by themselves they can give me knowledge. Nor can I justify a claim to know by appeal to intuition. Thus Rawls is not a traditional intuitionist and Hare is mistaken if he believes Rawls is subject to the objections raised against intuitionism in this sense.

But Rawls does appeal to some sort of intuition, and so we need to ask in what sense he is using the term. There is little

5Singer, p. 517.
indication in *A Theory of Justice*, but in his article "Outline of a Decision Procedure for Ethics" Rawls explains his use of "intuition":

By the term "intuitive" I do not mean the same as that expressed by the terms "impulsive" and "instinctive." An intuitive judgment may be consequent to a thorough inquiry into the facts of the case, and it may follow a series of reflections on the possible effects of different decisions, and even the application of a common sense rule, e.g., promises ought to be kept. What is required is that the judgment not be determined by a systematic and conscious use of ethical principles. It seems, then, that Rawls is using "intuition" in a purely psychological, as opposed to epistemological, sense to refer to a manner in which one comes to think or believe certain things. It does not by itself provide knowledge. Thus Rawls is an intuitionist in the sense that he believes there are moral judgments not determined by a systematic and conscious application of ethical principles and that these judgments play a role in the justification of ethical principles. Whether Rawls is to be criticized for his intuitionism depends on the role assigned to the intuitive judgments, i.e., on whether they are supposed to justify what they cannot.

To determine precisely the role of intuitive moral judgments in Rawls' justification procedure we need to distinguish them from two other types of moral judgment: considered judgments and judgments in reflective equilibrium. A considered moral judgment is an intuitive judgment made under conditions which in general are favorable for deliberation and judgment (ATOJ, pp. 47-48). Thus it must satisfy the criteria we would use in selecting considered judgments of any kind. It must be made impartially, freely, with confidence, after a careful inquiry into the relevant facts, etc. A moral judgment in

---

6 "Outline of a Decision Procedure for Ethics," *The Philosophical Review* 60 (1951), p. 183. Henceforth abbreviated in the text to ODPE. Rawls says that he still accepts "with some modifications" the point of view of this article (ATOJ, p. xi).
Reflective equilibrium is one made after reflection on the principles which govern one's considered moral judgment (ATOJ, pp. 20, 48-49). Unlike a considered judgment, a judgment in reflective equilibrium is not a kind of intuitive judgment since it is made with the conscious use of ethical principles. A reconsideration of one's considered judgment in light of ethical principles may cause one to reject that judgment. Or one may reject or revise the principles in favor of a deeply held considered judgment. In either case the resulting judgment is one in reflective equilibrium.

One of the restrictions imposed by Rawls on a considered judgment is that the judgment be stable, "that is, that there be evidence that at other times and at other places competent judges have rendered the same judgment on similar cases" (ODPE, p. 182). In other words, considered judgments are judgments on which there is a consensus. Rawls justifies placing this restriction on considered judgments on the basis that we cannot have confidence in a judgment concerning which there is disagreement by competent persons. The requirement of stability is important, for when Rawls appeals to intuition to justify some principle or judgment in A Theory of Justice it is always a considered intuitive judgment on which he is relying. That is, it is always a judgment on which there is a consensus. Thus we find "I assume there is a broad measure of agreement that," "commonly shared presumptions," "we are confident," "we think," and so forth.

If my interpretation of Rawls is correct, Rawls relies on considered judgments for justification purposes not because of their status as intuitive judgments, but because of their status as judgments on which there is a consensus. This interpretation is supported by Rawls' statement that justification must proceed "from what all parties to the discussion hold in common" (ATOJ, pp. 580-81). This means that Rawls has not assigned intuitive judgments a role in the justification procedure which they cannot perform, for he has assigned them no role as such. It is only when intuitive judgments are also judgments of consensus that they are appealed to. So if Rawls is subject to criticism it cannot be for his intuitionism. If criticism is to be made it must be for Rawls' reliance on consensus, a topic to be discussed later.
II

As Hare and Singer are aware, Rawls claims that his principles are objective:

We share a common standpoint along with others and do not make our judgments from a personal slant. Thus our moral principles and convictions are objective to the extent that they have been arrived at and tested by assuming this general standpoint and by assessing the arguments for them by the restrictions expressed by the conception of the original position. (ATOJ, pp. 516-17)

However, to Singer this constitutes "only the elimination of personal bias, and not real objectivity." Without presuming to know what is or is not "real objectivity," I think we can agree with Singer that there is an important sense in which objectivity is not identical to impartiality. For example, if in grading philosophy papers I do so impartially but without reference to criteria and merely by whim, I think it could correctly be said that my procedure for grading has not been an objective one.

However, I believe Rawls is thinking of more than just impartiality in the preceding passage. The principles of justice are objective not only because they are judged from a common standpoint, but also because the choice of principles is the outcome of a rational procedure, that of reflective equilibrium. Rawls is more explicit on the nature of objectivity in the early article:

The objectivity or subjectivity of moral knowledge turns... on the question: does there exist a reasonable method for validating and invalidating given or proposed moral rules and those decisions made on the basis of them? (ODPE 48)

According to Rawls we ought to require for objectivity in ethics no less nor no more than what we require for objectivity in science: that the propositions of the field be the result of "a reasonable and reliable method." And this is what he claims for the principles of his theory.

Singer, p. 494 fn.
Presumably Singer would find this characterization of objectivity still lacking. What bothers Singer is that "Rawls has left no room for any idea of validity that is independent of achieving reflective equilibrium." Rawls has made it true by definition that the principles which match our judgments in reflective equilibrium are valid. This is subjectivism "in the most important sense" for it implies that the validity of a theory is dependent on the judgments that people happen to make. Even if everyone's considered moral judgments were in agreement, this would at best produce intersubjective validity, not objective validity. Indeed, such a view is not only subjectivistic but also relativistic for it implies that if one society holds one set of considered judgments and another society holds a conflicting set, for each society there will be a valid set of principles and neither will be in error. According to Singer, the fact that reflective equilibrium implies this kind of relativism is something Rawls seems to accept.

A question needs to be raised here concerning Singer's use of "validity." If it means "objectivity," then it is true that Rawls has no idea of validity that is independent of achieving reflective equilibrium but only because he believes reflective equilibrium is the reasonable and reliable method the following of which constitutes objectivity. To be objective is simply to employ the rational procedure appropriate to the subject matter. In ethics this is the method of reflective equilibrium. The method of reflective equilibrium goes beyond mere intersubjectivity. Consider the possible situation where we all use the same but silly criteria for grading philosophy papers. Here there is intersubjective agreement but not objectivity because the procedure is not a rational one. In contrast, reflective equilibrium is a rational procedure. To ask for an idea of objectivity in ethics that is independent of following a rational procedure is like asking for an idea of objectivity in science that is independent of following the procedures of inductive logic.

If, on the other hand, Singer is using "validity" to mean "truth," then again it is true that Rawls has no idea of validity that is independent of achieving reflective equilibrium, but only

8 Ibid., p. 493.

9 Ibid., p. 501 fn.
because he believes there are no moral facts other than our considered judgments, i.e., no ethical characteristics such as rightness and wrongness which make moral judgments correct or incorrect. To say that a moral judgment is true is not to say that it has accurately described some moral reality, but to affirm that the judgment would be agreed to by rational persons using a reasonable procedure. Thus Singer is demanding from Rawls either an idea of objectivity that exceeds that required in science or an idea of moral truth that is at least questionable. In either case, Singer's charge of subjectivism against Rawls rests on a weak foundation.

Concerning the charge of relativism, we should first note that the only evidence Singer presents that Rawls accepts relativism is a reference in one of Rawls' footnotes to Quine's view of justification. Further, Rawls recognizes that intuitive judgments vary from society to society. But given the restriction of stability on considered judgments this does not imply relativism. There may still be one set of principles which explicates our considered judgments.

When we think of a successful explication as representing the invariant in the considered judgments of competent judges, then the variation of codes and customs is not decisive against the existence of such an explication. (ODPE, p. 70)

Rawls seems to think that as a matter of fact there are some considered judgments which competent judges of all societies would make and, therefore, there can be a unique set of valid principles of justice. Thus Rawls is not a relativist in the sense that he believes there are conflicting and equally valid principles.

Still, Rawls is committed to relativism in the sense that should there turn out to be nothing invariant in the judgments of competent judges then there may be conflicting sets of principles equally valid for different societies. However, this kind of relativism should disturb only those who believe there must be one set of principles applicable to all persons no matter what the contingencies of human existence. Rawls assumes that all justification must begin from what the parties to the discussion hold in common. Should there be no common agreement at all on ethical matters among human beings, ethical
relativism would seem a most plausible, even if unpleasant, theory to accept.

For Hare, Rawls' subjectivism arises not from a failure to provide a notion of validity independent of his justification procedure (Hare himself, I think, is guilty of this). Rather, Rawls is a subjectivist because he gives considered judgments "probative force" in that procedure. The truth (and objectivity) of a theory is tested against people's opinions rather than some firmer foundation, such as the logic of moral concepts. "He is making the answer to the question 'Am I right in what I say about moral questions?' depend on the answer to the question 'Do you and I agree in what we say?'."10 This, according to Hare, is subjectivism "in the narrowest and most old-fashioned sense." By making truth dependent on what people say, Rawls is committed to the view that thinking or saying something can make it so.11

I think Hare has much oversimplified Rawls' view. For example, Hare ignores the distinction between intuitive judgments and considered judgments. Theories are not tested against any view a person may happen to hold, only those views which are considered judgments. Further, Hare neglects to mention that theories are not directly tested against considered judgments. Rawls' procedure calls for a weighing process, balancing one's considered judgments against proposed principles, with neither having absolute authority. People's judgments, just as much as principles, are considered fallible. Thus it is not true that for Rawls my thinking or our thinking x is right makes x right. If x is right it is because x is the outcome of a reasonable and reliable justification procedure. That procedure ultimately appeals to our considered judgments and in that sense I suppose Rawls is a subjectivist. But to attribute to Rawls the view "thinking something can make it so" is surely misleading.

III

The important question, then, in evaluating Rawls' moral methodology is "To what extent should we rely on considered

10Hare, "Rawls' Theory," p. 145.

11Ibid., p. 146.
judgments as part of a reasonable justification procedure in ethics?" The extreme position is that of Singer who maintains "it would be best to forget all about our particular moral judgments, and start again from as near as we can get to self-evident moral axioms." Our particular intuitive judgments are untrustworthy, for they are likely the product of religious, sexual, and social views we now reject. At most, common sense morality can serve as a "useful check" on our intuitions of ultimate moral principles, never as a test of those principles. Recognition that an apparently self-evident principle conflicts with common sense morality should cause us to reconsider the genuineness of our intuition of the principle. However, the intuition may be genuine in spite of common sense morality.

Unfortunately Singer, like Hare, has neglected the distinction between particular intuitive judgments and considered judgments. The former are often influenced by prejudice and discarded systems of religion, custom, and morality, but the latter are selected under restrictions intended to eliminate those influences. The choice we face is not between self-evident axioms and particular intuitive judgments but between self-evident axioms and considered judgments. And, given the difficulties mentioned earlier which face any intuitionism of the type Singer is suggesting, a view which relies on considered judgments rather than self-evident axioms seems preferable.

Hare seems to grant considered judgments a larger role in the justification procedure than Singer. In "The Argument From Received Opinion" he says that while many common sense moral judgments are discarded when confronted with utilitarian arguments, there are also cases in which a consideration of the opinions of the ordinary person causes us to revise our moral principles or to recognize that we have misapplied them. This is because "many heads are better than one" and are therefore likely to perceive aspects of a question missed by a single person. So far this sounds somewhat like the method of reflective equilibrium. However, Hare departs from it by arguing

12Singer, p. 516.

that the way of resolving conflicts between principles and judgments, and the final arbiter in any case, is the analysis of the ordinary use of moral terms and the examination of the interests of the ordinary person. It is the appeal to the use of moral terms, and not to considered judgments, which has probative force. Hare agrees with Rawls that any justification must proceed from a consensus, but from a consensus on matters of logic and facts concerning the interests of the parties, not from a consensus on substantive moral questions. In­
deed, if you begin with a consensus at the level of considered moral judgments, it will hardly be surprising if you end up with a moral theory that supports all your shared intuitions.

I believe Rawls' response would be that the appeal to a consensus on substantive moral matters is inescapable (cf. ATOJ, p. 51). Consider how Hare handles the car-pushing example in "The Argument From Received Opinion." Anti-utilitarians present the case of six people pushing a stalled car over a hill when only five are necessary and argue that under certain conditions one of the six could justify on utilitarian grounds his merely pretending to push. But this would violate our con­sidered moral judgment that it is wrong to deceive others in this way. Consequently, utilitarianism is mistaken. In defense of the utilitarian, Hare argues that the considered judgment is correct, but not because there is a moral consensus. Rather, there are good utilitarian reasons for not deceiving the car-pushers. For the free-rider would be harming their interests by deceiving them. One can harm the interests of others even if they do not discover the deceit. Thus the considered judgment is determined to be correct through an examination of the interests affected by the action.

Notice, however, that in deciding the issue Hare makes a (factual?) claim concerning which there certainly is no consen­sus. Many would disagree that a person's interests are harmed by deceit even if he doesn't know about it ("What he doesn't know won't hurt him"). My point here is not that Hare is mis­taken in his claim, but that he cannot say he is appealing to a consensus on a factual matter. Since Hare agrees that

14Hare, "Rawls' Theory," p. 147.

justification must proceed from some consensus, he must appeal elsewhere for support. No doubt he would look in the direction of a definition of "interest" and provides such a definition in *Freedom and Reason*:

To have an interest is, crudely speaking, for there to be something which one wants, or is likely in the future to want, or which is (or is likely to be) a means necessary or sufficient for the attainment of something which one wants (or is likely to want).  

But is this the only way of defining "interest"? Isn't there room for disagreement here? For example, according to this definition it seems that to give heroin to an addict is in his interest. I'm not sure this isn't correct, but I doubt Hare would find a consensus on this. It seems then that in deciding whether the free rider's action is wrong the only consensus we can find is the agreement of our moral convictions. Hare's belief that he can escape from appeal to considered judgments is illusory.

Furthermore, even if a consensus on definitions and matters of fact was always available, it would not hold a privileged position in the justification procedure. Here Rawls presents Quinean considerations. Quine argues that scientific testing involves a challenge to a whole body of beliefs. Nothing is immune to revision. There may be occasions when we would reject a principle of logic rather than sacrifice certain factual beliefs based on our experience. For Rawls too, "justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view" (*ATOJ*, p. 579). In ethics those considerations include not only logical principles and factual beliefs, but also our considered moral judgments. We face moral problems with the whole system of beliefs. Though we are perhaps more willing to surrender certain considered judgments in the justification process, there is nothing we could not give up. "Definitions have no distinct status and stand or fall with the theory itself" (*ATOJ*, p. 51). Thus there may be ordinary uses of moral terms which we would reject or alter before sacrificing a deeply held moral conviction on which there is universal agreement.

---

Hare responds to the Quinean argument with the following:
Even Quine would hardly say that scientific theories as a whole are to be tested by seeing what people say when they have thought about them (it would have been a good thing for medieval flatearthers if they could be); but that is what Rawls is proposing for moral principles.\(^{17}\)

But again Hare has misinterpreted Rawls' view. Rawls does not hold that moral principles are to be tested by what people think about them. Considered moral judgments are not about moral principles. They are judgments about substantive moral questions at a level lower than that of principles. Quine would, I think, say that one test (though surely not a conclusive one) of theories of the shape of the earth is what competent persons say based on their experience.\(^{18}\) Rawls is simply suggesting the same for ethics: that one test of moral theories is what competent people say based on their experience.

Richard A. O'Neil
Philosophy Department
Emory & Henry College
Emory, VA 24327

\(^{17}\)Hare, "Rawls' Theory," p. 145.