

## **Ought-implies-can, the original position, and reflective equilibrium**

*Author:* Terence Rajivan Edward

*Abstract.* Are John Rawls's most noticeable methodological contributions, reflective equilibrium and the original position, consistent with each other? I draw attention to a worry that they stand in inconsistent relationships to the claim that ought implies can: it can only be the case that we ought to do something if we can do it.

*Draft Version:* Version 1 (27<sup>th</sup> September 2022).

*Claims of ought and their role*

*I pondered as I scanned this whole*

John Rawls is well-known in philosophy for recommending reflective equilibrium and the original position as methods for determining what would be a just set of institutional principles. Below I shall draw attention to a worry that there is an inconsistency between them involving the claim that ought implies can: that it can only be the case that we ought to do something if we can do it.

Let us begin with reflective equilibrium. A preliminary sketch of the method, not entirely reliable as a guide to Rawls proper,<sup>1</sup> will suffice for our purposes (or my purpose, to be precise). One proposes a set of general principles for institutions to implement. Then one tests this proposal against one's moral judgments about specific situations. Do they entail these judgments or not? For example, you own an apartment, but there is a service charge for maintaining the

---

<sup>1</sup> It overlooks Rawls's qualifications regarding which moral judgments one should set aside (1999: 42) and the option of entering more general moral judgments into the procedure as data.

building and the gardens. You have not paid it, because it can only be paid by bank transfer and bank transfer requires your debit card, which you cannot seem to find at the moment. The apartment management company contacts a relative of yours, whom you have disowned. But there are legal means of redress; you judge that contacting any relative, disowned or not, is the wrong approach. This is a moral judgment about a specific situation. If one's proposed principles do not entail most of such judgments, then one cannot maintain this exact proposal. It must be revised or abandoned. But Rawls argues that if it entails most of them, there is the option of abandoning judgments about specific situations which conflict with these principles. When principles entail judgments, then we have reached reflective equilibrium and both are justified.

Now we are disposed to make many judgments about many specific situations, some actual and some hypothetical. Some situations we have not even entertained yet, though if we did, we would make a certain judgment, or some of us would. Rawls thinks that the hypothetical is relevant. He writes:

There are, however, several interpretations of reflective equilibrium. For the notion varies depending upon whether one is to be presented with only those descriptions which more or less match one's existing judgments except for minor discrepancies, or whether one is to be presented with all possible descriptions to which one might plausibly conform one's judgments together with all relevant philosophical arguments for them... Clearly it is the second kind of reflective equilibrium that one is concerned with in moral philosophy. (1999: 43)

But it does not seem within human capacity to test proposed principles by considering each possible situation. This is Rawls's response:

To be sure, it is doubtful whether one can ever reach this state. For even if the idea of all possible descriptions and of all philosophically relevant arguments is well-defined (which is questionable), we cannot examine each of them. The most we can do is to study the conceptions of justice known to us through the tradition of moral philosophy and any further ones that occur to us, and then to consider these. This is pretty much what I shall do... (1999: 43)

It is not that easy to grasp what Rawls has in mind as a solution. Whatever it is, a natural interpretation is that he is committed to the following:

(A) We ought to reach reflective equilibrium based on testing proposed principles against all specific situations, actual and hypothetical, and our moral judgments about them. (See the final sentence of the first quotation.)

(B) We cannot consider all specific situations and our moral judgments about them. (See the opening two sentences of the second quotation.)

*Maintaining that combination requires rejecting the claim that ought implies can:* the claim that it can only be that we ought to do something if we can do it.

Let us turn now to the original position method. Rawls thinks that institutions should implement fair principles and that the method is to be used in determining which principles are fair. We are to imagine some self-interested individuals choosing from a menu of principles, on the assumption that the choice will be implemented by institutions in their society. But they are not like us in a crucial respect: they lack knowledge of various features of themselves, such as their sex, their class background, their talents, their plan of life, and more. The reason why they lack knowledge of these features is so as to prevent bias. A self-interested individual who has that knowledge will prefer principles tailored to their own case, such as a principle which allows

no one but members of their class to be a member of government. But they are not totally without knowledge. They have general social science knowledge. Rawls tells us, “They understand political affairs and the principles of economic theory.” (1999: 119) They are not to choose principles which cannot be implemented given their social science knowledge. The rationale for this seems to be: a principle tells us what we ought to do and we only ought to do that if we can do. *So the original position method, with the role it accords to social science knowledge, is justified by reference to ought implies can.*

Thus Rawls’s two renowned methods appear to be inconsistent: the one method involving a rejection of the claim that ought implies can and the other being justified on the basis of that claim. How then can reflective equilibrium and the original position be recommended together?

## **Reference**

Rawls, J. 1999 (revised edition). *A Theory of Justice*. Cambridge, Massachusetts: Belknap Press.