

REPRESENTATION AND OBLIGATION IN RAWLS'S SOCIAL CONTRACT THEORY

Simon Cushing
University of Southern California

There are two main roles for an account of justification in political philosophy. The first is that of political obligation, of settling the question of which society I am uniquely obligated to.¹ The second role for an account of justification is that of establishing whether or not a social system is a *just*, or more vaguely, a *legitimate* one. That the two justificatory roles are distinct can be seen by noting that even in a world of many perfectly just societies with perfectly legitimate laws, one would still presumably only be a citizen of one of those countries. These roles are not clearly distinguished in the classic social contract theories of Locke and Rousseau in particular, but there the contract fills both roles, as I will show. Rawls's hypothetical social contract seems designed only to address the second role, but Rawls does in fact explicitly tackle the issue of political obligation in *A Theory of Justice*,² where he tries to remedy the flaws of the classic social contract *obligation* account with a *duty-based* account. In this paper I shall assess the criticisms of his account offered by A. John Simmons in his book *Moral Principles and Political Obligations*³ and a possible response to the criticisms that Samuel Freeman has implicitly offered in his work on Rawlsian social contracts. I conclude, however, that this response rests on an equivocation in *Theory* and thus fails.

1. The Founding Contract

Consider the following model, intended as a very rough sketch of a literal reading of Locke's social compact: Art, Billie, Charlie and Dinah decide to form a club. They openly discuss rules of behaviour for members of the club, the scope of the rules, positions of authority in the club, rules for election into those positions, term limits, etc. Having come to a consensus on all points, the four of them jointly agree to found the club, which currently comprises the four of them. Call the agreement they reach the Founding Contract, and the sketch in general the Club Model.⁴

The Founding Contract provides an *obligation* account of the club members' moral requirement to follow the rules: Art et. al. are all bound to follow the rules, *because* they agreed to them. Consider Ella, for example, who sat in on the discussion of rules, contributed many valuable suggestions which the others all agreed made the rules better, but declines to join in the

actual act of agreement. Nobody would say, I think, that Ella is a member of the club or bound to follow its rules. The actual act of agreeing to be a member (whether it take the form of swearing an oath, shaking hands or whatever agreed form of acknowledgment) is required to found obligation.

What about legitimacy? The source of legitimacy of the rules *for those members* is the fact that they openly discussed them and came to an agreement on them (you might say this was the bargaining stage of the Founding Contract, without which it is not a contract but simply an agreement).

Thus if the social contract approaches our rough sketch, then different aspects of the same contract serve both to legitimize the social ordering and bind members to it. Granted, the Club Model is a crude sketch, but it does serve to highlight the strengths and weaknesses of the idea of the founding contract as found in Locke, for example. As an account of political obligation it has these strengths. An actual agreement by an actual citizen to abide by the laws of her society, provided it takes place in the right, non-coerced circumstances (admittedly a big proviso) is intuitively a plausible sufficient condition for a moral obligation. Furthermore, the fact that it is a social contract means that the justificatory strategy respects the *autonomy* and *individuality* of the members of the club, which is what particularly appeals about the use of a social contract in political philosophy, in contrast, in particular, to various versions of utilitarianism.

However, the major flaws are of course the following: first, it is by no means clear that founders of actual countries agreed in this way. Second, even if they did, their consent does not bind us. And third, although we could be bound, like Ella, by an act of *joining* consent, in real instances such a clear, intentional sign of agreement, in conditions of full knowledge of alternatives, is almost never given. And the flaws in the alternative notion of *tacit* consent (say, on the grounds that we live or own property within the borders of a country) are well-documented.⁵ Thus, the Founding Contract fails to be *general* enough to account for the obligation of at least the large majority of citizens, who live peacefully within a society and indeed if asked would regard themselves as citizens. It is this lack of generality that prompts Rawls's alternative, duty-based account, to which we now turn.

2. Rawls's Duty-Based Account

In *A Theory of Justice*, Rawls states:

From the standpoint of justice as fairness, a fundamental natural duty is the duty of justice. This duty requires us to support and to comply with just institutions that exist and *apply to us*... Thus if the basic structure of society is just, or as just as it is reasonable to expect in the circumstances, everyone

has a natural duty to do his part in the existing scheme. Each is bound to these institutions independent of his voluntary acts, performative or otherwise. [p. 115, my emphasis]

Rawls's account of political obligation⁶ is in terms of a natural duty that we have to comply with just institutions. Thus it is only obliquely contractarian, in that the contract is a way to test whether or not an institution is just, its justice being a necessary condition for the application of the duty to comply. Let me sketch a second model to illustrate roughly how Rawls's social contract contrasts with the Club Model.

In this case, Art, Billie, Charlie and Dinah are particularly concerned that the rules of their club (or, to be more precise, the general principles that will frame the rules) be *fair*, but are worried that if they bargain over the rules themselves, their personalities and relationships will distort the principles in favour of the stubbornest or most forceful personality among them.⁷ So, instead of working out the rules of their club themselves, they each hire an agent, and get the agents together jointly to choose from a set of conceptions that our prospective club members themselves provide as exhausting the most reasonable alternatives to produce principles to order their rules. The agents are not informed about any particular features of their clients, but each consultant is told to represent *only* her client's interests. Assuming that Rawls is correct, and that the perfect rationality of the agents (the parties of his original position) combined with the veil of ignorance that divides the agents from their clients (which models perfect equality) ensures that the political conception the agents choose is fair (and by extension, *just*), it seems that on Rawls's picture the four *clients* are obligated to follow the principles chosen by their *agents* because of each client's natural duty of justice. Call this model the Consultant Model. I will postpone further comment until I have laid out Simmons's criticisms.

Note that in his description of the natural duty of justice Rawls includes the stipulation that the just institutions *apply to us*. This seems a reasonable enough stipulation, for it preserves the distinction between political obligation and an account of justice, without which we could not explain why we are citizens of one just state over another.⁸ Simmons seizes on just this clause, however, as the basis of his criticism that Rawls's duty account is inadequate to ground a bond of citizenship. He analyses the possible senses in which an institution could be said to "apply to" an individual, and suggests that they fall into three rough categories: the "weak" sense, the "territorial" sense, and the "strong" sense.

Simmons argues, and I think he is right, that in *Theory* Rawls assumes the "territorial" sense of application but that this does not intuitively

