Doing history in the original position

Author: Terence Rajivan Edward

Abstract. An objection to John Rawls's original position is that it faces a problem of inconsistent features: the individuals in this hypothetical situation are not supposed to know where they are in history, but they have knowledge of general social science, from which they can infer at which point in time they are. In this paper, I consider two solutions. One of these solutions depends on extending a solution to another well-known objection: that readers cannot imagine lacking the knowledge that these individuals lack.

What a fruitful idea John Rawls's original position is! But of course there are objections to it. One well-known objection is "I cannot even imagine this situation Rawls is asking us to imagine." We are to imagine some self-interested individuals agreeing on the rules for society. In order to ensure a fair agreement, they lack knowledge of various features of themselves, such as their talents, their sex, their class, and their conception of a good life. Otherwise if seven of the individuals have talent A and three have talent B, for example, then the seven will prefer rules which favour individuals with talent A over talent B, such as that only those with this talent have the right to freedom of movement. In this paper, I focus on another objection to the original position, which I call "The social science history problem." But before even introducing that objection, I am going to consider the "I cannot imagine this" problem in more detail, because a solution to it is relevant.

1

The "I cannot imagine this" problem

The condition of not knowing that Rawls places these individuals in is called the veil of ignorance. More than one philosopher raise the objection that we cannot imagine being an individual behind the veil, in order to think as they would – to determine which rules these people lacking biasing factors would prefer (Sandel 1984: 86, 90; Okin 1997: 68-70). At present I am not convinced by this objection. We can distinguish between at least two interpretations of what we are being asked to do, two interpretations of what it means to imagine being behind the veil of ignorance. The objection does not apply on both interpretations.

The what it's like interpretation. On the first interpretation, to imagine being behind the veil is to carry out a task similar to what novelists sometimes do. A novelist might well try to imagine the point of view of someone who is quite different to themselves, such as an old fisherman on a remote island. A more unusual example is a science-fiction novelist who tries to imagine the point of view of a person whose mind has somehow been tampered with and who now lacks knowledge of their talents, their sex, their class, conception of what a good life would be, and other potential sources of bias. How does such a person interact with others? How do they feel in various situations? On the first interpretation, Rawls is asking us to perform a task similar to the one this science-fiction novelist attempts. How do they feel about certain rules? "They worry about this," "They feel safe with that," and so on. We can call this "the what-it's-like interpretation," because we are asked to imagine what it is like to be a person who lacks the knowledge Rawls places behind the veil.

"How can I even do that?" someone scrawls irritably in the margins. Their thinking, more fully, seems to be as follows. The method Rawls recommends must

meet two conditions:

(a) It must not only be available to people suffering from a strange madness, but also to some people within the range of the psychologically normal.

(b) It must be rational for us to believe that some people within this range can do more than use the method; they can use it to lead to results that we can rely on.

But it is not rational to expect the psychologically normal fellow to be reliable when imagining the point of view of people with all this ignorance, so that we can be confident that what is described is their point of view. There are different ways of filling in the details of why, the obvious starting point being that the thinking of a normal fellow is too much influenced by awareness of their own talents, sex, and other features.[1]

The premise-by-premise interpretation. On the second interpretation, it is not especially useful to compare the thought experiment to what novelists sometimes try to do. The comparison is potentially misleading. Rawls is simply asking us to not include some information when constructing arguments for why individuals in the original position would prefer a certain option, more formally to not include premises that refer to this information (1999: 103). What one has to do then is identify the premises of the argument, check that none of the banned premises are present, and then evaluate the argument almost as one normally would. Are the premises acceptable and is the reasoning to the conclusion valid? I say "almost" because objections when conducting the thought experiment must also not appeal to the banned information.

Rawls himself encourages this interpretation, when he tells us "one or more persons can at any time enter this position, or perhaps better, simulate the

---

[1] "They can only imagine a person who prefers satisfying their specific interests"; "They can only fleetingly imagine the appropriate person, before their interests intervene"; etc.

deliberations of this hypothetical situation, simply by reasoning in accordance with the appropriate restrictions." (1999: 119) We can do the relevant checks! By the way, in everyday life people sometimes make arguments which do not refer to much or all of the information that Rawls places behind the veil of ignorance. For example, a worker does 10 hours work. A boss who pays the worker £10 per hour calculates that, given the contract, the worker should be paid £100. The argument for this conclusion does not make reference to the worker's talents, sex, class, conception of a good life, or even whether they like the worker or not.

The social science history problem

The use of "the" here may be misleading if there is another problem that deserves to be called a social science history problem, but the well-known one is as follows. Amongst the things which individuals in the original position do not know, along with their talents, class and the other things listed above, is at which point in history they are at. This is so the rules chosen are not tailored for the conditions faced by one particular generation and can function across generations. But amongst the things which they do know is social science[2] of a general nature. Of their society, "They do not know its economic or political situation," writes Rawls (1999: 118), but "They understand political affairs and the principles of economic theory." (1999: 119). They are supposed to use their general social science knowledge when selecting rules. Rules which this science entails cannot be stably implemented should not be chosen. But the social science knowledge will enable them to work out where they are in history.

There are different ways of explaining why. One explanation is that the

---

[2] A preliminary definition: "social science" refers to sciences focused on the study of human societies.

findings of social science include findings about when certain knowledge of society appears. For example, there is a finding which says that within a year after the invention of money and not before, knowledge of society which features proposition X appears, e.g. the proposition that barter economies are inefficient.[3] The social science knowledge of individuals in the original position features proposition X. So now they can infer that they are in a society after the invention of money. Let us call this "the law explanation," or the law version of the problem: it posits relationships of social science law between knowledge of society and where a society is in history, that it is a law that this knowledge appears at this point. I later consider another explanation, after introducing two solutions.

The banned inferences solution. This is an extension of the solution presented earlier. To use the original position, what one does is present an argument in premises for choosing certain rules. The argument can include premises which are social science propositions but what it cannot include, apart from the banned premises, is any inferences to the period in which the arguer is in, even if these inferences are valid given the premises. We do not just ban certain premises, which refer to the information that Rawls prohibits use of, but also certain valid inferences.

The adapted social science solution. We give individuals an adapted social science from which they cannot tell where in history they are. Sometimes social scientists develop ways of doing social science for peculiar conditions, such as when they lack historical knowledge (Radcliffe-Brown 1952: 3). Similarly, we give them a social science that is specially adapted for the conditions of the original position, meaning that, even if they always infer validly, they cannot work out which period they are in from this science. One might compare the original position to a model

---

[3] I took this to be a general finding, or hypothetical general finding if it is false, because it identifies a relationship that potentially obtains within societies not in contact with each other. Rawls suggests a fancy alternative: general findings are formulated without rigid designators (2001: 69).

aeroplane, which comes with its own equipment for putting it together and warnings such as "Only use the glue provided."

Dimming. At this stage, I wish to consider another explanation for why there is a social science history problem, suggested by D.D. Raphael, and how the adapted solution might cope with that explanation. Raphael does not focus on the findings of social science, rather the justifications (1974: 122). If individuals in the original position are to count as having knowledge, they must not only grasp these findings, he says, but also the justifications for them. These justifications involve dated experiences: "the following was observed on such and such a date…" So individuals are going to know that they exist at or after that date and are going to have some information about what occurred then and before.

Raphael's version[4] of the social science history problem does not suppose that there are laws connecting the social science of a society to its place in history[5]; it is the justifications for whatever social science they rely on that give away where they are in history. A response to this version is that individuals in the original position do not know the justifications for the social science, only the conclusions – the findings. I think that Rawls does not want to include dated justifications. He writes, "the course of history is closed to them; they have no information about how often society has taken this or that form, or which kinds of societies presently exist." (1999: 176) However, the justifications may well be relevant for deciding between two sets of rules. Both cohere with nine out of ten social science findings, say, but the finding that one set fails to cohere with has the support of less empirical data than the finding that the other set fails to cohere with. I suppose they could have access to findings along with ratings indicating justificatory strength but no details.

---

[4] It is not clear to me that this is exactly what he had in mind, but I think any differences are small.

[5] It depends on an empiricist account of knowledge, so perhaps there is a kind of law that this amount of knowledge could not have been there on day one. I have not considered the solution "It's all innate."

Another response, quite a strange one, is to give them social science findings with justifications but nevertheless it is a social science that has been selected or developed so that it is difficult for them to date more precisely when they are making the agreement. For example, the justifications only refer to empirical information that was available one hundred years ago. Perhaps they are still going to know that it is 1920 or after, but no more than this. The social science has been adapted so that the light it shines on their place in history is dim. I shall take "dimming" to be a way of pursuing the adapted social science solution, although it does give some historical knowledge. An advantage of dimming is that it helps with another problem: some of Rawls's social science commitments are from his time and no longer relied upon much. That is not transparently a problem if one is dimming. (It seems to me that problems for Rawls's philosophy are usually addressed one-by-one, when solutions have to be evaluated by how they help with multiple problems. Owing to the number and variety of problems, this allows for surprising results.)

Reality versus dimmed social science

The problem we are considering is how to make features of the original position consistent: the lack of knowledge of where one is in history and yet the knowledge of general social science. But when we consider the adapted social science solution, we run into another inconsistency. I assume individuals must take into account social science knowledge because Rawls accepts a kind of ought implies can.[6] You only ought to implement rules chosen in the original position if you can stably do so and the social science knowledge helps determine which options can be stably implemented. But if individuals in the original position are working with

---

[6] Rawls also accords the knowledge a role in achieving two other qualities, which are distinct, to the browser anyway: the principles lead to a desire to adhere to them (one can desire the impossible) and are not a great strain to commit to (some doable projects are a great strain). See 1999: 153-154.

specially adapted social science, to "dim" the knowledge of when they exist, then how can it perform this function? It may be that our best social science gives other verdicts on what is stable and what not – what can be stably implemented.

An answer is that the outcome of the original position procedure is not the final word on which rules should be implemented. "It is difficult to develop a coherent fair procedure for rule selection and the adapted social science solution is the best we can do to solve a problem; but the solution is not something perfect. Owing to imperfections further inquiries must be made after the procedure. If our best social science tells us that the option preferred by original position individuals is stable, then we implement it; but if it tells us that this option is unstable, then we reject it."

But what if our best social science does not give such clear-cut verdicts, rather we learn from it that an option that individuals in the original position rank as second best is a little more likely to be stable than their preferred option? The larger body of empirical data that we have, compared to individuals in the original position, supports this likelihood assessment.

Some people will say, "We just implement the original position option, because it has been chosen in fair conditions and the grounds for overturning that choice are too weak." Call these people "No Reversers." Other people will say, "Let us choose what individuals in the original position regard as second best. There are no grounds for saying that we ought to implement their top preference, because the social science that they have access to is not the best for determining what we can and cannot stably do." Call these people "Reversers." They are going to worry that contemporary social science is being ignored, in favour of company glue social science! Acceptance of the original position as a starting point does not protect against this divide. I quite like the Reversers' position: "Original position – yes; social

science history problem – yes; adapted social science – yes; dimming – yes," and still they are not going to the end of the journey; "Reverse the first and second choice preferences for reality please!" (See also Raz 1986: 128.) In the long-term, the pleasant first choice would seem to require the sensitivity of a great poet to get the taxation right.

References

Okin, S. 1987. Justice and Gender. *Philosophy and Public Affairs* 16: 42-72.

Radcliffe-Brown, A.R. 1952. Introduction to *Structure and Function in Primitive Society.* Glencoe, Illinois: The Free Press.

Raphael, D.D. 1974. Critical Notice: *A Theory of Justice* by John Rawls. *Mind* 83: 118-127.

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

Rawls, J (edited by Kelly, E.). 2001. *Justice as Fairness: A Restatement.* Cambridge, Massachusetts: Belknap Press.

Raz, J. 1986. *The Morality of Freedom.* Oxford: Clarendon Press.

Sandel, M. 1984. The Procedural Republic and the Unencumbered Self. *Political Theory* 12: 81-96.