# Collective Rationality: A Dilemma for Democrats with a Solution through Deliberation? Comment on Philip Pettit

Natalie Gold

In his paper, Pettit presents a paradox that may arise when a group must vote on an interconnected set of propositions. Even though every individual within the group may be perfectly consistent in her own views, or individually rational, the propositions that the group would accept as the result of a majority vote on each individual proposition are inconsistent, so the group is collectively irrational. Pettit distinguishes between two ways a group may deal with this dilemma: the rationality-first approach, where the institution is committed to some mechanism for prioritising one of the propositions despite the votes of the individuals within it, and the responsiveness-first approach, where, although re-votes may be taken, the institution has no formal means for prioritising propositions when individually rational preferences aggregate to be collectively irrational. Given the argument presented in the paper, you will have to forgive me if I say that, whilst I agree with the conclusion, that we should take a rationality-first approach, I debate the propositions on which that conclusion is based. In my response I will explain why I think that, for all practical purposes, the rationality-first approach is inevitable for democratic institutions, regardless of supporting arguments that deliberative democrats might supply. This implies that the real argument is not whether we should achieve collective rationality but how, so I shall go on to examine whether arguments from deliberation can shed any light on this issue. Finally, I will say something about putting deliberation more directly into this type of decision-making model.

Regardless of contestability and conversability, as Pettit himself points out, the rationality-first approach is effectively enforced on most deliberative institutions. There is an analogy between the institutional and individual agents that may help to make the reason for this clear. An individual has a belief system consisting of multiple propositions which she endorses. Then in decision-making situations her choices are based on these. For an individual, the belief system is primary, in the sense that the beliefs exist prior to the need to make decisions or to act and a person may hold beliefs that she will never act on. In contrast, most institutional deliberation is primarily about deciding what actions to take and the endorsement of propositions happens purely as a stage in this process. Although the action to be taken may depend on the endorsement of certain premises by either the individuals within the group or the group as a whole, inevitably the key question is whether

to convict the defendant or which policy to implement. And, unlike the individual case, the decision is probably only taken once, or at least fairly infrequently, so that if there are any inconsistencies in the propositions endorsed, agenda manipulation is more likely to be a problem than the possibility of cycles which is so often discussed in the individual case.

To continue the comparison, whilst an individual may take a responsiveness-first approach and endorse a set of propositions which provide conflicting reasons for action, when the time comes to act these inconsistencies must somehow be resolved. Otherwise the individual faces paralysing indecision with potentially disastrous consequences, as did Buridan's ass. So some reason is found for prioritising certain propositions in the belief system. Of course, an individual could chose to do nothing, but this often amounts to doing something in the sense of promoting a particular outcome. Or another response could be for the individual to chose to randomise between possible actions but, in practice, it is the case that individuals do not like to randomise. Experimental evidence shows that humans are very bad at randomising when asked explicitly to do so[1] but, more importantly, when the issues at stake are important we tend to think that 'I flipped a coin' or 'I threw a dice' is not an acceptable justification for action. We believe that actions should be based on reasons relevant to the case in hand. Unlike an individual agent, an institution does not usually have the luxury of a period where it can endorse conflicting propositions before having to choose an action. The endorsement of propositions happens synchronically with the making of decisions so that, unless we are willing to advocate that our institutions randomise or do nothing when faced with this type of paradox, the rationality-first approach will effectively trump the responsiveness-first approach. This conclusion does not depend for its force on any arguments from deliberation.

The question we are left with then, is which propositions should take priority when inconsistencies arise. Although it might seem that the group itself could decide in each instance, there is reason to think that guidelines should be chosen in advance. If both procedural and substantive reasons are decided at the same time, then there is the possibility of strategic voting by agents on the first question, of which proposition takes precedence, in order to ensure their favored outcome on the second question, of what the policy should be.

As Pettit points out, a deliberative group may be able to achieve consistency by modifying its judgement on any of the three propositions. However, I have nothing to say about this case but instead will focus on the cases where, as in the judicial example, the

---

[1] Smith (1999).

propositions fall naturally into premises and conclusions. I think that this actually covers more cases than is apparent at first sight. If one grants that people tend to form their judgements about complex propositions from their judgements about simple propositions then a large class of problems, including those involving conjunctions and disjunctions, will have this form, with the simple propositions as the premises and the complex ones as the conclusions. The analogy with the individual case might then suggest that the collective judgements on the premises should drive the collective judgements on the conclusion. That is also the approach that Pettit's republican deliberative democratic conversability and contestability arguments support. I do think that, where the group is such that it is natural to think of the collective as having a single agency, there is a presumption for resolving the dilemma in this premise-driven manner. (So this might cover 'the party thinks...' or 'the government says…' but I am not sure that it includes the judges in the original example.) However, there are other considerations that drive in the opposite direction and I think that there are occasions on which these have force. In order to elaborate on these, I shall draw on work done elsewhere, to which Philip Pettit has been a chief contributor.

As is mentioned in the paper, there is a line of argument which, following the Marquis de Condorcet, advocates democratic procedures because of their increased chance of 'getting it right'. This argument favors a conclusion-driven approach to the paradox, with individual judgements on the conclusion being aggregated to get the collective judgement. The simple intuition is that voting on the premises makes the collective more likely to be correct about the premises and therefore right for the right reasons. But, since the conclusions are a complex combination of the premises, there are other ways of reaching the correct conclusion, though for the wrong reasons. For instance, if the truth is $P \vee Q$, then one could get to the correct conclusion by believing $P \wedge \neg Q$, $Q \wedge \neg P$ or $P \wedge Q$ but only one of these combinations of premises is true. The conclusion-driven approach maximises the collective's chance of reaching the right judgement, albeit for the wrong reasons.[2] The argument by Pettit is that, if a republican deliberative democrat cares about conversability and contestability, then she should care about being right for the right reasons and therefore endorse the premise-driven approach.[3] However, I think that the analogy to the individual case shows why there may sometimes be exceptions to this rule.

---

[2] Pettit and Rabinowicz (2001).

[3] Pettit (2001).

When we consider beliefs it is as, if not more, important to endorse a proposition for the right reasons as it is to endorse the correct proposition. But, as I have said earlier, collective decision-making is generally with regard to proposed actions, not abstract belief systems. With actions, it is possible that it is better to do the right thing for the wrong reasons than to increase one's chance of having the right reasons at the expense of one's chance of doing the right thing. In other words, when actions are involved, consequences arguably matter as much as reasons. A concrete example of relevance to deliberative institutions is the criminal trial. In our society we have a preference that the innocent go free. If we take the judicial example in the paper but change it to a murder trial and let P = the defendant killed the victim, Q = the defendant had intent to kill the victim then the verdict of murder depends on the truth of the conjunction $P \wedge Q$. In this case we may prefer to maximise our chances of reaching the correct verdict. It is more important simply to be right than to be right for the right reasons.

Another type of occasion when deliberative institutions may be more concerned about conclusions than premises is when making certain political decisions. In a multi-cultural society we need to find a basis acceptable to all on which we can live together. We search for a Rawlsian ‘overlapping consensus’ on our public institutions even if citizens have differing underlying conceptions of the good. For instance, consider changing the example in that Pettit’s paper to the questions of whether we should allow parents to opt out of state schools and educate their children privately. Let P = parents have a right to educate their children privately, Q = we should educate our children within their religious faith. In this case we use a disjunction rule and if $P \vee Q$ is true then we should allow private schooling. Now there might plausibly be three different groups in our society. Libertarians believe $P \wedge \neg Q$, the religious believe $Q \wedge \neg P$ and Liberals believe $\neg P \wedge \neg Q$. If we aggregated premises then we would have a majority for $\neg P$ and a majority for $\neg Q$, so we would not allow people to opt out of the state system. However, if we aggregate conclusions then we would get a majority in favor of $P \vee Q$, so we would allow private schools. In this case we may prefer to aggregate individual judgements on conclusions precisely because we believe that the job of public institutions is not to judge the premises. We are not looking to enforce the ‘right’ reasons on those we disagree with but to get enough agreement on conclusions about how we structure our society so that its institutions will be acceptable to differing groups within it. Within the political arena, following Rawls, some liberals explicitly do not judge the individual conceptions of the good underlying our shared public intuitions. For them, it is more important to agree on conclusions than to have the right reasons.

Even a deliberative democrat with a presumption in favour of having the right reasons may think that there are occasions when it is more important to either be right about or to agree on the conclusions than to be right about or to agree on reasons. But, anyway, it seems to me that all the preceding arguments could as easily be made by a Rawlsian who believes in public justification without any reference to deliberative democracy. So far, I have said a lot about democracy but very little about deliberation.

Finally, let me make some comments about where deliberation might fit into this model. The paradoxes presented are all aggregation problems. In other words they may be faced in any democratic decision-making situation, not just those involving deliberation. Deliberative democrats may have particular reasons to argue that the dilemma should be solved in one way rather than another but the model itself does nothing to capture the deliberation process. Models cannot be expected to capture everything, by their nature they simplify situations in order to clarify key points. One key feature of deliberation, which is often used as an argument in its favour and which is missing from the model, is that deliberation changes peoples' preferences. During deliberation some types of preferences are ruled out and people begin to come to a consensus. These aggregation models assume that voters' opinions are independent of each other. If they have spent time deliberating, then one might hope or expect that this is not the case. Adjusting the model to reflect simple correlation of voter preferences does not affect the mathematical results. However, another way in which deliberation may change preferences is that, rather than correlating opinions on particular issues, it causes voters to structure the argument and the issues at stake in a common manner. It has been formally proved that if deliberation induces sufficient preference structuration then deliberative democrats will avoid such doctrinal paradoxes.[4] Such arguments provide a resolution of the dilemma that is truly unique to deliberative democrats.

**References**

List, C. (2001) *Mission Impossible? The Problem of Democratic Aggregation in the Face of Arrow's Theorem*, D.Phil Thesis, University of Oxford.

Pettit, P. (2001) "Deliberative Democracy and the Discursive Dilemma", *Philosophical Issues* (supplement 1 of *Nous* 35) 11, 268-95.

[4] List (2001).

Pettit, P. and W. Rabinowicz (2001) “The Jury Theorem and the Discursive Dilemma”, Appendix to Pettit 2001, 295-99.

Smith, J. (1999) “How Hard is it to Randomise? Theory and Experimental Results”, *Santa Fe Institute working paper* 99-07-048.