

## Book Review

*Group Agency: The Possibility, Design, and Status of Corporate Agents*, by Christian List and Philip Pettit. Oxford: Oxford University Press, 2011 Pp. ix + 238. H/b £25.00.

In this fascinating and enjoyable book, List and Pettit argue that it is logically possible that there exist *group agents*, that is, singular composites, constituted by human agents, which possess, process, and act upon representational and motivational attitudes (beliefs and desires), in much the same way as human agents. (Corporations, governments, committees etc., are held up as possible candidates).

List and Pettit further argue that such agents, should they exist, are fit to be judged responsible and held responsible, and to be the possessors of rights and duties, and hence are *persons* (at least on a Hobbist-Lockean ‘forensic’ use of the term). They also detail how, and to what extent, it is possible for group agents to be *well-designed*, or as one might say, fit for their members’ purposes. In particular, we are told about the sorts of functions from individual to group propositional attitudes that, if implemented by a group’s organizational structure, are conducive to its being a reliable indicator and tracker of truth, and to its not constraining its members’ freedoms — even in the face of the various non-group minded motives that these members are likely to have.

The chapters on organizational design are among the book’s most engrossing. Many of the claims here summarize technical work on interpersonal judgement and preference aggregation (much of it by List), and whilst we are given only informal versions of the key theorems, and sketches of their proofs, well-chosen examples convey the main ideas. Some of the results here will bring the pleasure of recognition to those, not least those employed by universities, with experience of the psychopathologies of committee life. For example, it is shown that, even when all group members have a strong preference for truth-tracking, there can be incentives for them to misrepresent their judgements to the rest of the group. In particular, when every member believes that consensus is building around a falsehood, a love of truth may be trumped by a mild preference for conformism, even when both are universally held, since, under majority voting, the casting of the sole dissenting vote

achieves nothing but nonconformism. As List and Pettit put it, in game-theoretic terms, for individuals with slight conforming preferences, the situation in which all vote for  $p$ , and the one in which all vote against  $p$  are both Nash equilibria, regardless of anyone's individual judgement on  $p$ .

In the remainder of this review, I focus on the authors' central *modal* claim, that of the logical possibility of group agents. Readers steeped in the debate between ontological individualists and holists may be puzzled by the point of this claim. For individualists, who argue that group agents do not exist (typically on the grounds that we do not need to postulate their existence to explain any phenomena) are likely to be unmoved by an argument for a claim that does not contradict their view — it is not as if individualists have tended to argue that group agents are *impossible*, only that, like, say, ectoplasm and fairies, they are unnecessary posits. And holists, who argue that group agents exist (typically on the grounds that we need to postulate their existence to explain certain phenomena) are likely to be unimpressed by an argument for what is a mere corollary of their view.

Of course, if List and Pettit do not mean to intervene in the ontological debate, then it is no criticism of them that they do so ineffectually. But they *do* take a clear stand on this debate. At page seventy-five and page four they make, in terms, the ontological claim that group agents actually exist. (The claim is also tacitly presupposed by many discussions in the text, as when contrasts are drawn between the models of group organizational structures offered by aggregation functions, and what are regularly called 'real world' group agents). One wonders then, what the *point* of the weaker, but much more heavily advertised, modal claim can be, given the authors' commitments on the ontological issue. Why ink into your subtitle, as well as your opening paragraph, a relatively weak and dialectically ineffectual claim (the modal claim) while more quietly asserting and presupposing a stronger one (the ontological claim)? The puzzle is compounded by the fact that, while much is said in support of the weak claim, very little is said in support of the stronger one. Indeed, the brief argument given (at p. 75) for the ontological claim seems to draw entirely on considerations supporting the modal claim:

groups of individuals can in principle count as agents. Let a collection of individuals form and act on a single, robustly rational body of attitudes ... and it will be an agent; let it function or perform in the role of an agent, and it will indeed be an agent ... it is possible for collections of individuals to coordinate their individual contributions so as to achieve this level of functioning. Hence group agents exist.

It would be *very* uncharitable to suppose that the authors had offered an argument for a weak claim, quietly asserted a stronger one, and encouraged their less careful readers to assume that the latter followed from the former. Whilst I do not suppose that this is the authors' strategy, they might have done more to explain what it is.

Here is what I *think* they are up to. The authors think that there is defeasible evidence for the ontological claim, namely, perhaps inter alia, the apparent veracity of our everyday and social scientific attributions of acts and attitudes to groups. (For, they seem to imply in their introduction, the hypothesis that such talk is in error, metaphorical, or a mere *façon de parler*, is unmotivated). The problem with the defeasible evidence, however, is that it may be defeated, for there is a prima facie case for the impossibility of group agents (I will shortly try to reconstruct this case). It is for this reason that the authors set about undermining this case. The hope, I take it, is that they thereby establish the modal claim, and, thereby, vindicate the defeasible evidence for the ontological claim.

The prima facie case for the impossibility of group agents proceeds from two requirements that the authors take to be non-negotiable. First, supervenience: for any actual or possible group agent, its attitudes and actions supervene on those of its individual members (p. 66). For List and Pettit, this amounts to a denial of what they take to be the historically dominant realist view of the relation between group agents and their members: emergentism. Second, rationality: any actual or possible agent of any sort, is, at least by and large, rational (p. 24). For the authors, this seems to be a conceptual necessity. That these two requirements, when applied to putative group agents, might seem to pull in different directions, can be seen from the following matrix, which depicts some of the beliefs held by the three members of a committee:

	p	q	p & q
Member 1	True	True	True
Member 2	True	False	False
Member 3	False	True	False

Each committee member has a consistent set of beliefs regarding the propositions  $p$ ,  $q$ , and  $p \& q$ , so nothing in the matrix rules out their meeting the rationality requirement on agency. But does the committee itself meet this requirement? It may seem that it does not. For by supervenience, any attitudes that *it* possesses are determined by the contributions of its members. The most obvious way in which they could be so determined is by simple majority voting. But if the committee has beliefs, determined by majority voting, regarding the propositions  $p$ ,  $q$ , and  $p \& q$ , then it inconsistently believes  $p$ ,  $q$ , and  $\neg(p \& q)$ , so does not satisfy the rationality requirement on agency.

This is not yet an argument for the *impossibility* of group agents. What the authors report, however, is a proof that, on a precisified formulation of the two requirements, conflict between them is inevitable. The precisified notion of group-individual supervenience restricts its focus in two ways: *first*, to

supervenience of group *attitudes* on member *attitudes* (i.e. actions are ignored (p. 66)); *second* (although this is not made explicit, and sits uneasily with the authors' stipulation (p. 30) that a 'group' is a collection of people with 'an identity that can survive changes of membership') to groups with a fixed, unchanging membership. Precisified supervenience is also necessary rather than contingent in the sense that a group supervenes on its members only if *for any possible combination* of member attitudes on the propositions towards which the group has attitudes, the member attitudes determine the group attitudes (p. 67). As for the authors' precisified notion of rationality, the authors give it a term of art, 'robust rationality'. And for them (p. 47ff., p. 67), a group agent's beliefs are robustly rational just if they are (i) consistent, in that for no inconsistent propositions does the agent believe each one, and (ii) complete in that, for each proposition-negation pair among some logically connected propositions (ones at least as numerous and connected as  $p$ ,  $q$ , one or other of  $p \& q$ ,  $p \vee q$ , and  $p \rightarrow q$ , and the negations of these) the group believes either one of the pair, and each group member believes either one of the pair too. Generalising beyond belief, a group agent's *attitudes* are robustly rational just if the above is true both of their beliefs and *mutatis mutandis* their desires also.

Given these stipulations, three Arrovian theses of increasing generality may be proved, report the authors (p. 67ff.). First, that robustly rational group attitudes cannot supervene by means of a *majoritarian* function, applied, for each proposition, to the member attitudes to it, to yield the group attitude to it. Second, assuming an embargo on the 'degenerate' function of 'dictatorship', by which a group inherits the attitudes of a single member, that robustly rational group attitudes cannot supervene by means of *any* function (e.g. a qualified majority, or unanimity function) applied, once again, for each proposition, to the member attitudes to it, to yield the group attitude to it. Third, again assuming no 'dictatorship', and that unanimous member attitudes to a proposition ensure the same group attitude to it, that robustly rational group attitudes cannot supervene by means of *various* functions (qualified majority here, simple majority there, etc.) by, for each proposition, a function being applied, to the member attitudes to it, to yield the group attitude to it.

It is these proofs that constitute the *prima facie* case for the impossibility of group agents. It is only *prima facie*, for they do not show that robustly rational group attitudes could not supervene on member attitudes. But they do show that they could not *neatly* supervene on member attitudes.

The case, however, may be undermined. For (the authors report) it is provable that robustly rational group attitudes can supervene on member attitudes. It is just that, for them to do so, they must *holistically* supervene, in the sense that at least some group attitude to a proposition is determined, not by members' attitudes towards *that* proposition, but by their attitudes to

other (e.g. logically related) propositions (p. 69). For example, robust rationality may be restored to the committee of our matrix simply by its coming to believe that  $p \ \& \ q$  and ceasing to believe that  $\neg(p \ \& \ q)$ . It is just that if this happens, the group's attitude towards *those* propositions is settled, not by its members' attitudes to them, but by their attitudes to other, logically related propositions ( $p$ , and  $q$ ).

This, of course, is ingenious, but I advise any reader to ask themselves, not merely whether an Arroviaan sledgehammer has been used to crack a nut, but whether this was a nut that needed cracking in the first place. Is any philosophical problem that anyone ever had, or should have had, answered by the authors' reported proof of the consistency of supervenience and rationality as *they precisify them*? For since supervenience is precisified as *attitude-on-attitude* (and hence as prescinding from individual and group actions), as *static* (prescinding from membership change) and as *necessary* (obtaining for any possible combination of member attitudes), and since rationality is precisified as *robust* rationality (i.e. roughly, as consistency and completeness) it is at the very least an open question whether, thus precisified, these are fair requirements on group agents. I must say that they strike me as blatantly unfair ones. Nor do the authors attempt to argue that they are fair requirements. Furthermore, even granting these requirements, it is not, I should say, clear that the realist about group agents should ever have expected them *non-holistically* to supervene on their members, or to have been perplexed, or felt their realism threatened, by the impossibility of this. Nor is it clear, and nor do the authors explain, why of all the conceivable defeaters of the defeasible case for realism about group agency, it is this impossibility alone that deserves such extensive and technical treatment. I am reminded of Kripke's animadversion that

Logical investigations can obviously be a useful tool for philosophy. They must, however, be informed by a sensitivity to the philosophical significance of the formalism and by a generous admixture of common sense, as well as a thorough understanding both of the basic concepts and of the technical details of the formal material used. It should not be supposed that the formalism can grind out philosophical results in a manner beyond the capacity of ordinary philosophical reasoning. There is no mathematical substitute for philosophy. (Saul Kripke, 'Is There a Problem About Substitutional Quantification?' *Truth and Meaning*, ed. G. Evans and J. McDowell, Oxford: OUP, 1976, p. 416)

The authors' result that only *holistic* supervenience is consistent with robust rationality has, they believe, an interesting corollary. It is that whilst a group agent's attitudes, would, by supervenience, be *determined* by those of their members, they would also, by the holism of this supervenience, be *autonomous* of them too, in the sense that they would not easily be discernible in, or recoverable from, any patterns exhibited by the individual member attitudes. On this basis, List and Pettit suggest, group agents, would be both

*unmysterious* — their existence and nature fixed by those of their members — but also *non-redundant* ontic postulates — real additions to the furniture of the world — not mere neat abstractions from agents and their attitudes. (In this, List and Pettit say, they differ from, say, weights, which are mere neat abstractions from certain truths about objects).

But notice those ‘would’s. Even assuming no mismatch between supervenience and rationality *as fair requirements on group agents*, and supervenience and rationality *as precisified by the authors*, the proof of supervenience-rationality consistency, does not, by itself, show that there are any unmysterious and non-redundant group agents, only that, were there any group agents, they would be unmysterious and non-redundant. Still, the foregoing suggests an alternative reconstruction of List and Pettit’s overall strategy. For perhaps they mean to proceed from the worry that the defeasible evidence for the ontological claim, that is, the apparent veracity of our everyday and social scientific attributions of acts and attitudes to groups, *would* be defeated, *were* such talk easily ‘paraphrasable away’, as talk of individual agents. Their response is then to show that, since group agents would have to be autonomous, in the sense identified above, there could be no easy paraphrase, so the evidence is undefeated. Hence, ‘the lack of an easy translation of group-level attitudes into individual level attitudes requires us to recognize the existence of group agents’ (p. 5).

If *this* is the authors’ strategy, it rests on a fallacy. Grant, for argument’s sake, that, were there group agents, truths about their attitudes would not admit of easy paraphrase into truths about their members’ attitudes. It does not follow that our *actual* everyday and social scientific attributions of attitudes to groups are not thus paraphrasable. Besides, as is shown by ordinary object eliminativists such as Peter van Inwagen (1990, *Material Beings*, Cornell University Press), Keith Hossack (2000, ‘Plurals and Complexes’, *British Journal for the Philosophy of Science*, 51: pp. 411–43), and Cian Dorr (2002, *The Simplicity of Everything*, Princeton University Ph.D. dissertation, accessible at: <<http://users.ox.ac.uk/~sfop0257/papers/SimplicityOfEverything.pdf>>), paraphrases of sentences that apply predicates to singular terms for putative composite objects, into sentences committed only to the putative parts of those putative composites, can be a lot easier than one might have expected — one need only read them as applications of *collective* predicates to *plural* terms. So we should doubt that there are no easy individualist translations, even if there *are* autonomous groups.

All my criticisms notwithstanding, there is much to ponder, and much to learn from, in this lively work. The authors acknowledge their debts to writers who have reached similar conclusions, by comparable argumentative means, such as Carol Rovane. It’s a shame that Susan Hurley’s *Natural Reasons*, which in several ways anticipates elements of the authors’ strategy, receives no mention.

*Department of Philosophy*  
*University of Manchester*  
*Oxford Road*  
*Manchester M13 9PL*  
*UK*  
doi:10.1093/mind/fzs073

THOMAS H. SMITH