Being and Becoming
in the Theory of Group Agency

Leo Townsend

University of Cape Town
South Africa
leo@tenderscan.co.za

Abstract

This paper explores a bootstrapping puzzle which appears to afflict Philip Pettit's theory of group agency. Pettit claims that the corporate persons recognised by his theory come about when a set of individuals 'gets its act together' by undertaking to reason at the collective level. But this is puzzling, because it is hard to see how the step such a collective must take to *become* a group agent – the collectivisation of reason – can be taken without them already *being* an agent. I explore this puzzle by recounting Pettit's account of the emergence of group agents. According to Pettit this process has two stages: a first stage in which a collective incurs the distinctive pressure exemplified by the Doctrinal Paradox, and a second in which the collective responds to that pressure by instituting decision-making mechanisms designed to secure collective rationality. After arguing that this second, response stage in Pettit's account is not coherent, I conclude with the tentative suggestion that the personhood of groups should be seen as depending not only on the efforts of group members but also on the recognitive attitudes of other persons in a wider discursive community.

Introduction

Over the past ten years Philip Pettit has developed a novel form of realism about group agents.[1] Drawing on formal results obtained in the field of judgement aggregation, Pettit argues that suitably designed social groups should be recognised as 'corporate persons'[2] with 'minds of their own'.[3] Such conclusions have been advanced before, but what is especially striking about Pettit's theory is that it resists capture in either of the two dominant traditions through which they have previously been promulgated. The first of these traditions is the 'authorisation theory' associated with Hobbes, according to which a group agent is formed when a collective authorises a privileged sub-group to speak and act for the whole. Pettit contends that, unlike his own view, this conception of group agents is only a 'redundant realism', since 'everything the recognition of such a group agent entails is already expressible in individual-level language'.[4] But Pettit also wishes to distance his theory from the other venerable tradition of theorising

---

[1] Pettit (2002; 2003; 2007; 2009), List & Pettit (2006; 2011). Throughout this paper I refer to the theory of group agency as Pettit's theory, even though it has been largely worked out in collaborative publications (most regularly with Christian List). This is because Pettit is the common thread in all the publications I am focused on, and, judging by his individually-authored work, it is most likely he who is responsible for the sections of the co-authored work which are of primary interest to me here.

[2] Pettit (2002), List & Pettit (2011)

[3] Pettit (2003)

[4] List & Pettit (2011: 8)

about group agents, which he calls 'animation theory'. Animation theory, which Pettit traces to the influence of Hegel, holds that 'group agents are emergent entities over and above the individuals who compose them'[5], formed when a transcendent force 'animates them and gives rise to a single centre of agency'.[6] On account of his commitment to 'methodological individualism', Pettit insists that this sort of view is metaphysically objectionable and intolerably mysterious. The view he defends thus seeks to be both non-redundant and unmysterious, and my hope for this paper is that it casts some doubt on whether Pettit successfully achieves this.

Pettit claims that the corporate persons recognised by his theory of group agency are 'made, not born'[7], and this raises the question of just who it is who makes them. His original treatment of the Doctrinal Paradox and subsequent forays into organisational design made in collaborative work with Christian List[8] suggest that it is the group itself that somehow 'gets its act together', transforming itself from a collection of individuals into an autonomous rational agent.[9] But this creates something of a bootstrapping puzzle for his account, for it is hard to see just how the step that groups must take in order to *become* agents – the 'collectivisation of reason' – can be taken without the group in question already *being* an agent.

My aim is thus to see whether Pettit succeeds in providing a non-redundant, unmysterious and *non-circular* account of the process whereby a collective incorporates into an agent or person.[10] In Section 1, I explore what Pettit considers to be the first stage of this process, in which a collective incurs a certain kind of *pressure* – the pressure to incorporate, to become a group agent. For Pettit, this stage of the process is exemplified by the much-discussed Doctrinal Paradox of judgement aggregation, and so the question I pursue in this first section is whether the paradox can have any traction for a collective that is not already incorporated. In Section 2, I turn to the second stage in Pettit's account, in which the collective responds to the pressure to incorporate by undertaking to exercise reason at the collective level. Pettit thinks there are a variety of decision-making procedures which might accomplish this 'collectivisation of reason', and the question I ask about these is whether they too might actually presuppose an already-incorporated group agent. In the end I think that Pettit fails to dissipate the bootstrapping problem bruited above, and, in Section 3, I diagnose this failure as the product of a tendency to overlook the role that inter-subjective attitudes in a broader discursive community might play in instituting a group's status as an intentional subject or corporate person.

1   The Pressure to Incorporate

The lesson of the Doctrinal Paradox, and Pettit and List's[11] feted impossibility theorem, is that two requirements we may wish to place on the formation of collective attitudes are in fact in tension with each other.[12] The first is that the collectively-held attitudes ought to be sensitive to the attitudes of the collective's members; the second is that the attitudes collectively

---

[5]Ibid., p. 9

[6]Ibid., p. 9

[7]Pettit (2007: 495)

[8]Pettit (2002); List & Pettit (2002; 2011)

[9]For the sake of clarity I use the term 'collective' here to denote an un-incorporated but potential group agent, and 'group' for an already incorporated group agent.

[10]Worries over the possible circularity of accounts of collective intentionality have been explored before, notably by Tollefsen (2002b) who targets the views of Tuomela (1995) and Gilbert (1996). As far as I know Pettit's theory of group agency has not been subjected to similar critique.

[11]List & Pettit (2001)

[12]Though Pettit's recent work (esp. List & Pettit 2011) focuses predominantly on the impossibility theorem, I concentrate here on the Paradox. I take it that Pettit's reason for this recent shift in focus is that the impossibility theorem somewhat more *precise* than the Paradox: it sharply delimits the extent of the difficulty for attitude aggregation (as not simply a problem for majoritarian aggregation functions, but all such functions), and it allows Pettit to identify exactly

endorsed should comprise a set that would be rational to hold. Now Pettit's key claim is that certain kinds of collectives will face an overwhelming pressure to give priority to this second requirement, even though doing this may lead them to violate the first requirement – they may end up espousing views which many (even all) of their members individually reject. To make sense of Pettit's account of how group agents emerge we must first make sense of this 'pressure' to exercise reason at the collective level.

1.1   Paradox and Pseudo-paradox

The Doctrinal Paradox is usually illustrated with the following toy jurisprudential example. A defendant is being sued for damages arising from an alleged breach of a contract. Legal doctrine dictates that the verdict of the court – as to whether or not the defendant is liable – must depend on just two issues: (1) whether a contract existed prohibiting the defendant from performing a certain action, and (2) whether or not the defendant performed that action. Such a case will be turn out to be paradoxical if the views of the judges have a certain kind of structure, such as the following:[13]

Judge	Issue 1: Contract?	Issue 2: Breach?	Outcome: Liability?
A	Y	N	N
B	N	Y	N
C	Y	Y	Y
Majority	Y	Y	N

The reason this profile is paradoxical is that the outcome of the case depends on the decision-making protocol adopted. If the court allows a majority of outcome votes to determine the overall outcome – sometimes called the 'conclusion-based approach' – the defendant is not liable; if they apply legal doctrine to the majority-favoured issue votes – the 'premise-based approach' – then the defendant is liable.

Pettit has devoted considerable effort to generalising the scope and significance of the Doctrinal Paradox in this form – showing how it can arise for a great variety of groups in a great variety of circumstances, and that it can arise even without the mediating influence of 'doctrine'.[14] But for all his generalising efforts, Pettit has, to my mind, neglected the task of *circumscribing* the difficulty confronted by groups in paradoxical scenarios. For it is possible – and, I think, instructive – to generate profiles which *resemble* paradoxical cases but in which, intuitively, there is no real problem for the group. As we shall see, reflection on such 'pseudo-paradoxes' can make it tempting to think that the pressure to collectivise reason will only be felt by groups which are already incorporated.

Imagine that a television station wishes to know whether viewers would like the Olympic diving event to replace one of the two programmes – Game of Thrones (GT) or Mad Men (MM) – ordinarily screened between 7pm and 9pm on a certain night. Upon interviewing a (very) small sample audience, the following profile is generated:

---

what the various options for overcoming that difficulty are (i.e, the different ways of relaxing the conditions which generate the impossibility theorem). From the perspective of my discussion here, however, these considerations are both moot: I am concerned only with the general tension between collective rationality and individual responsiveness (which the Doctrinal Paradox reveals vividly), and I am happy to admit that collective rationality can be secured by the various decision-making procedures Pettit proposes. Moreover, since the Paradox gets us to reflect on the way in which a particular kind of collective, in the course of its discursive activity, might encountering and respond to this 'discursive dilemma', it is arguably more amenable to the question I am concerned with here – the question of the *process* through which group agents emerge.

[13]Kornhauser & Sager (1993: 11)

[14]See esp. Pettit (2003)

Person	Replace (GT or MM) with diving?	Drop GT?	Drop MM?
A	Preferred	Dispreferred	Preferred
B	Preferred	Preferred	Dispreferred
C	Dispreferred	Dispreferred	Dispreferred
Majority	Preferred	Dispreferred	Dispreferred

This profile may well present a problem for the TV station, but it does not obviously amount to a paradox for the TV-watchers. A majority of the individuals interviewed preferred that regular programming be interrupted in order for the diving to be screened, but most did not want Game of Thrones to be dropped, and most did not want Mad Men to be dropped. Is 'the majority' confused here, even though each individual is quite clear and consistent? No, because the apparent inconsistency involved in 'what the majority wants' vanishes when we observe that on each issue the majority preference is constituted by a unique pair of individual preferences. Although it is inconsistent to wish for either Game of Thrones or Mad Men to be dropped, but neither Game of Thrones nor Mad Men to be dropped, the rational requirement that such inconsistency be avoided only applies to unitary entities. Since the majority featured in each column is, strictly speaking, a *different* majority each time, the bottom line of this profile reflects disagreement amongst the people interviewed rather than inconsistency.

Why can we not then say, to return to the standard example of the Doctrinal Paradox, that there is no failure of rationality in the court's verdicts derived from majoritarian voting on all issues? This would mean that the court judges that (a) a contract existed, that (b) a breach occurred, that (c) if a contract existed and a breach occurred then the defendant is liable[15], and that (d) the defendant is not liable. This sounds irrational only because we have surreptitiously slipped from talk of what 'a majority of judges' thinks to talk of what 'the court' thinks; if we remember that the view of the court is only an arithmetic function of the individual judges' views then this appearance of irrationality disappears. In other words, to say that there is a failure of rationality appears to beg the question of the court's status as an entity in its own right – one which is subject to certain constraints of rationality.[16] But if paradox only arises for groups that are already rational entities in their own rights, then it seems it could not provide the impetus for the emergence of such entities.

1.2   Purpose and Paradox

Clearly, Pettit needs a principled way of distinguishing paradox from pseudo-paradox, and a way of deeming the bottom line inconsistency in the judicial profile pernicious without presupposing that the panel of judges is already incorporated. How might this be done? Perhaps by drawing on the idea that certain social collectives are *purposive*.

The sample group of TV-watchers is what Pettit calls a 'mere collection'[17]: they have no common purpose or even any knowledge of one another's contribution to the interview process in which they each participate. By contrast, the court is a 'purposive collectivity'[18] required to *act* in designated ways – imposing penalties, awarding damages, sentencing, etc. – which uphold legal doctrine. Because the court's judgements are 'near the coal-face'[19] of

---

[15]This is based on the assumption that the judges agree on the relevant legal doctrine.

[16]Such question-begging is, arguably, exhibited in Pettit's tendency to describe *the set of views* which issue-by-issue majoritarian voting would deliver as 'irrational' instead of simply inconsistent (see esp. Pettit 2002, 2003; List & Pettit 2002).

[17]List & Pettit (2011: 31-2)

[18]Pettit (2002, 2003)

[19]Pettit (2009: 77)

their action they cannot, on pain of jeopardising the court's agential capacities, be inconsistent: 'Let an agent try to act on inconsistent representations or motivations… and there will be a straightforward breakdown: actions will be supported that cannot be realised'.[20]

This distinction between purposive collectivities and mere collections seemingly helps Pettit to make non-circular sense of the pressure to incorporate. A collective might be recognised as purposive – it might have certain goals which it needs to pursue – and because of this, *as a response*, it ensures that its collective views are rational. According to Pettit, standards of rationality are 'nothing more or less than desiderata of agency: standards such that agents will generally do better as agents by robustly satisfying them'.[21] In order to be able to effectively *act* in fulfilment of its purpose, a purposive collectivity must take steps to ensure that its collectively-held attitudes are, at a minimum, mutually consistent.[22] Thus it is the panel of judges' purposiveness which distinguishes them from the TV-watchers, and which ensures that they find themselves susceptible to the distinctive pressure exemplified by the Doctrinal Paradox. In short, the fulfilment of purpose requires agency, and agency requires rationality – where, crucially, what is 'required' is not pre-required but can be provided responsively.

One might still wish to press the original bootstrapping objection here, by claiming that identifying something as a purposive collectivity must presuppose that it is a rational agent, since a purpose could not be fulfilled by a collective that was not already a rational agent. But this line of objection is mistaken: it elides the distinction between potentiality and actuality, overlooking the fact that treating something as a purposive collectivity can itself serve to instil or awaken the very rational-agential capacities required for the fulfilment of its purposes. This kind of bootstrapping, according to which certain capacities are *brought about* by being taken to be *already there*, is often thought to underpin the developmental processes in individuals. For example, some theorists[23] have claimed individuals come to be trustworthy only by being trusted – that is, by being treated as though they were already trustworthy. Similarly, it has been claimed that actual fitness to be held responsible for one's actions results from a process of 'responsibilization'[24] – of being treated as though one were fit to be held to responsible. Such bootstrapping seems to serve a powerful purpose in the individual case and it should not be ruled out in the case of social groupings.

A better line of objection targets the purported link between purposiveness and the requirement of agency. An agent, according to the conventional picture Pettit adopts, is something with beliefs and desires, and the capacity to intervene in the world for the fulfilment of its desires in line with its beliefs[25]. But collectives are composed of individuals who are themselves agents, so why can't the individuals comprising the collective co-ordinate their actions so as to serve the purposes of the collective? Purposiveness may require agency but it is not clear at this point why it requires the collective's own agency.

To answer these questions we need to attend more carefully to the notion of purpose involved in the idea of a purposive collectivity. Once again we may find it instructive to reflect on the differences between the TV-watchers and the panel of judges. Why are the TV-watchers not counted as purposive? Certainly they are interviewed for a purpose. The

---

[20]List & Pettit (2011: 25).

[21]Pettit (2007: 497)

[22]According Pettit, to be rational one must form attitudes which the evidence favours ('attitude-to-evidence' rationality), one must not have inconsistencies in one's body of attitudes ('attitude-to-attitude rationality'), and one must tend to act for the realisation of one's preferences in line with one's beliefs ('attitude-to-action rationality').

[23]E.g., McGeer (2002); Pettit (1995)

[24]Garland (2001)

[25]Pettit (2002, 2007, 2009); List & Pettit (2011)

crucial difference between them and the panel of judges seems to be that it is not *their* purpose but the TV-station's which their stated preferences are rung in to serve. (To see this, imagine that the interviewees pooled their resources and bought the TV-station; then the pattern of their judgements on the scheduling issue would be *their* problem, and the profile would be paradoxical.) By contrast, the panel of judges seems to have a purpose of its own – something which the individuals on the panel collectively pursue – so perhaps it is this proprietary purpose which prompts certain collectives to form their own centres of agency.

What then does it mean for a collective to have a purpose of its own? Presumably it means that there exists some goal that the collective strives to achieve or some role that it seeks to perform. But one thing it cannot mean, if Pettit's account of the emergence of group agents is to be coherent, is that the collective has a kind of *attitude* of its own – something like a group intention. For if the purpose of a purposive collectivity were tied to a group-level intention then it would seem that we would already have an incorporated group agent – something which is already an intentional subject in its own right – and so the having of a collective purpose could not be integral to the process of becoming such an agent.

Pettit sidesteps this concern by appropriating a key concept from the literature on collective intentionality – the concept of 'joint intention'.[26] According to Pettit, a joint intention can be distinguished from a group intention since it is a configuration only of individual attitudes, such that:[27]

(1) each of the individuals in the collective intends that they together promote some purpose,

(2) each of the individuals intend to play their part in fulfilling that purpose,

(3) each of the individuals forms these intentions (1&2) because they believe others have formed similar intentions, and,

(4) this is all a matter of common awareness.

The notion of joint intention appears to provide Pettit with a route out of the bootstrapping worry, at least insofar as the worry concerns the pressure to collectivise reason. The first clause (1) tells us that, where there is joint intention, each of the members of a collective intend that they – the collective – jointly serve some purpose. *Intending-that* (as opposed to *intending-to*) is conceived by Pettit as a matter of the individuals each favouring a state of affairs over which they have at least partial control.[28] This ensures that each of the members has an individual interest in taking steps to preclude the sorts of inconsistencies in collective-level judgements which can (as the Doctrinal Paradox reveals) result from perfect individual responsiveness. This is because, as Pettit tells us, inconsistencies cannot provide a sound basis for the fulfilment of the collective's purpose. The pressure felt by certain collectives to collectivise reason is precisely the realisation that their purposes cannot be served if inconsistencies in group-level judgements are allowed to arise. But when commitment to these purposes is understood in terms of joint intention it is an entirely individual affair – there is no group intentional subject presupposed.

2 The Response to the Pressure

Perhaps then a conceptual wedge can be driven between the notion of members of a collective severally sharing a purpose and the notion of a group with its own purpose. If so, then the fact that members of a collective share a purpose does not imply that the collective is already

---

[26]Bratman (1999); Tuomela (1995); Gilbert (2001)

[27]List & Pettit (2011: 33)

[28]As Pettit, drawing on Bratman's (1999: Ch 8) seminal statement of the concept, puts it, 'I intend that something happen only if... I want it to happen and I am in a position to do something about it' (Pettit & Schweikard 2006: 23).

incorporated – and hence it is possible that its purposiveness, understood in terms of joint intention, might call for incorporation without presupposing it. This would mean that we can make sense of the pressure to incorporate in a way that does not make it unnecessary (i.e., such that the only possible subjects of that pressure are already incorporated). But this does not yet mean that the coherence of Pettit's account of the emergence of group agents is secured, because it is not clear at this point how a collective could possibly *respond* to that pressure without already being a group agent.

We might worry that what is required as a response to the pressure exemplified by the Doctrinal Paradox is the collective's own response, where this is understood as some sort of group action. But such a response could not of course be part of the account of the emergence of group agency, since it would quite clearly require the collective to already be an agent. So what is needed is a member-level story – one which doesn't pre-require but might still be able to institute a group agent. Pettit seeks to provide such a story by endorsing a second notion prevalent within the collective intentionality literature – the notion of *joint action*, which he distinguishes from group agency.

Pettit claims that several individuals can co-participate in a joint action on the basis of their joint intention without constituting a group agent.[29] But although it need not, joint action can, according to Pettit, give rise to group agents: a collection of individuals can jointly act in certain specific ways so as to transform themselves into a group agent.[30] This is the second, response stage in Pettit's account of how group agents come into existence, and its tenability depends on whether the steps that need to be taken to transform a collection of individuals into a group agent are indeed candidates for joint action. In this section I tackle this issue by considering first whether collectivising rationality is *itself* a possible joint action, and second (because I think it is not) whether collective rationality is something which could be secured within a group indirectly, *via* the members' joint action.

2.1 Collectivising Rationality as Joint Action

What steps must be taken for a collective to transform itself into a group agent? We know from Pettit's discussion of the Doctrinal Paradox that he thinks they need to *collectivise rationality*, but is this a potential joint action? More particularly, can each individual member *intend that* they together collectivise rationality? This might seem a peculiar question, especially in light of the fact that we are not here questioning whether collectivising rationality is something that groups can do (our question is about *who* does it, and, specifically, whether it can only be done by a group agent). But if groups do collectivise rationality, and we agree with Pettit and List's[31] supervenience thesis – that the attitudes and actions of the group supervene only on the attitudes and actions of individual members – then it seems the collectivisation of rationality must be something which individual members can intend. Intending-that meant favouring a certain state of affairs – in this case, that the collective collectivises rationality – and committing to playing one's part in that state being realised. But since, by the supervenience thesis, individuals cannot help but be involved in anything the collective does (including becoming rational at the collective level), it seems that individuals can certainly so intend, and hence that the collectivisation of rationality can be a joint action.

We should, however, be careful here, for it is not obvious that simply being a part of something amounts to 'playing one's part'. Consider the following example. Imagine that,

---

[29]Pettit & Schweikard (2006: 19)

[30]Ibid., p. 33

[31]List & Pettit (2006, 2011)

for whatever reason, I wish to discover *completely by accident* that my partner is having an affair (perhaps I wish to separate, but think it is only legitimate for victims of adultery to instigate these things, and I also don't want to snoop or pry). Since it could *happen* that I discover completely by accident that my partner is having an affair, it seems plausible that it is a happening I could wish for. But is it something I could intend? Certainly I could not *intend to* make this discovery since what is discovered is not wholly within my personal sphere of influence (i.e., some part of it is in *another's* sphere of influence). Could I *intend that* I accidentally discover my partner's infidelity? I do not think so, even though it is plausibly a state of affairs I could favour, and one, were it to be realised, which would necessarily involve me. The reason I cannot intend it (in the sense of intending-that) is that the way in which I would have to be involved is not as an agent.[32] Since what has to happen for the wish to be realised has to happen *accidentally*, I cannot intend to actively play any role in bringing it about; playing a role would, precisely, defeat the fulfilment of that wish.

I think that something similar goes on when we try to conceive of the members of a collective intending that they together collectivise rationality. Collectivising rationality, as the Doctrinal Paradox teaches us, is in direct conflict with individual responsiveness. What this means is that, although this collectivisation must necessarily involve individuals, it does not cast them in the sort of role they can intend beforehand to perform. To commit to the collectivisation of rationality is to commit to being less than an individual agent in the exercise of one's group role; in acting and thinking *for* the group one is not thereby performing as an autonomous agent, but as part of the group's subpersonal machinery. When one intends, one takes up an attitude towards a potential future action of one's own (this can be part of a joint action), but what is required in the collectivisation of rationality is not the performance of one's own action, but the performance of someone else's action – the group's.[33] This means that the collectivisation of rationality is not a potential joint action; it is a singular process whereby a new centre of rationality and agency is formed, and this process is not one in which group members can intelligibly intend to participate.

2.2   Collectivising Rationality via Joint Action

Perhaps Pettit is able to sidestep this concern too, since he does not quite say that collectivising rationality is itself the joint action through which groups emerge. Nonetheless he does believe that joint action can effect the collectivisation of rationality, and so bring about group agents:

> 'First, the members act jointly to set up certain common goals and to set up a
> procedure for identifying further goals on later occasions. Second, the members
> act jointly to set up a body of judgements for rationally guiding action in support
> of those goals, and procedures for rationally developing those judgements further
> as occasion demands. And third, they act jointly to identify those who shall act
> on any occasion in pursuit of the goals'[34]

What is the difference between this and jointly performing the action of collectivising rationality? The key difference seems to lie in the introduction of 'procedures', by which is meant more or less explicit rules to guide the internal machinations of the group. If by joint action these procedures can be inculcated, and if action conforming to these procedures will count as collectively rational, then it appears that joint action can, in the end, bring about group agents (though that action is not itself collectivising rationality).

---

[32]In Pettit's terms, I am not 'in a position to do something about it'.

[33]Cf. Rovane (2004)

[34]Pettit & Schweikard (2006: 33)

But what will these all-important procedures be? Pettit considers the virtues of a number of functionally explicit organisational procedures, including the 'sequential priority rule' in which logically interconnected issues are placed in a sequence and issues only go to a vote so long as they are not already determined by their logical relation to prior, already-decided issues. In the end Pettit thinks that though they guarantee consistency in group-level judgement, the rigidity of functionally explicit procedures such as the sequential priority rule may cause groups endorsing them to violate some other demands of rationality: [35]

> 'When I realize that some propositions that I believe entail a further proposition, the rational response may well be to reject one of the previously accepted propositions rather than to endorse the proposition entailed. Those are the undisputed lessons of any coherence-based methodology and the group that operates under a sequential priority rule, or under any variant, will be unable to abide by them'[36]

So if functionally explicit procedures are unable to secure collective rationality, what sort of organisational mechanisms does Pettit think groups should adopt in order to transform themselves into group intentional subjects? To answer this we need to attend to a distinction Pettit makes between 'reason' and 'rationality'. As we saw above, rational standards, for Pettit, are nothing other than the desiderata of agency; rational creatures are those who tend to believe what the evidence favours, whose store of beliefs and preferences are largely consistent, and who tend to act in line with their intentions. They are veracious, consistent and enkratic because this is the best way to be an agent. But none of this requires a rational creature to attend to the propositional contents of her attitudes and form meta-propositional attitudes about them – asking, for example, whether this belief is *true*, whether that preference *implies* another, whether these judgements are mutually *consistent*, and so on. According to Pettit, simple rational agents can get along adequately without being able to attend in this higher-order way to their thought; their rationality – the extent to which they meet the standards of rationality laid out above – is 'secured by their make-up or design'.[37] But sophisticated rational agents, who have the capacity to form attitudes about their attitudes, can go beyond this, reinforcing their rationality by means of *reason*:

> 'To be able to reason... is to be able to conduct an intentional activity that is designed – and perhaps explicitly intended – to raise the chance of satisfying [rational] desiderata'[38]

Pettit thinks that the robust rationality of groups will only be secured if the group can succeed in *reasoning*, and he thinks that there is at least one sort of organisational procedure which can achieve this: the straw vote. A straw vote is simply a non-binding vote, through which a group's judgement can be registered without being concretised; the views thus reached can always be revised later if the group so decides. A straw vote can be used to temper the inflexibility of the likes of the sequential priority rule, thus allowing groups to be more sensitive

---

[35]Functionally explicit decision-making procedures like the sequential priority rule and the distributed premise-based approach are also, arguably, susceptible to a being-becoming circularity worry similar to the kind I explore here. Determining an issue sequence or assigning extra weight to the views of designated experts within the group are themselves important decisions which the group faces and which may themselves depend on prior issues. If so, then in order to avoid an endless regress, the group must seemingly be already able to decide the issue of sequencing / expertise assignment before these procedures can be deployed. But then it looks like the purported mechanisms of collectivising rationality actually presuppose a group that is already collectively rational (and able to make decisions). This concern has been developed by O'Madagain (2012), but since the problem does not appear to afflict the procedure which Pettit truly favours – the straw vote – I am not pursuing it here.

[36]Pettit (2007: 511)

[37]Ibid., p. 495

[38]Ibid., p. 499

to what rationality (not just consistency) requires. For example, a group might vote on each proposition in a sequenced set, and, if their voting generates an inconsistency, they 'consider all the different possible ways in which previously formed attitudes or the new attitude could be revised so as to restore consistency… [and] take a vote… on which of the possible revisions to make'.[39]

When groups deploy a straw vote procedure they are clearly doing what Pettit would call reasoning. This is because they are concerned with the consistency of the judgements they have defeasibly adjudicated – standing ready, should inconsistency be discovered, to make the appropriate revisions *at any point within the overall body of judgment* (depending on what the evidence favours). In this way the straw vote allows the collective to be more robustly rational, since the collective is required to reflect on the overall rationality of their collective judgements.

Once again, this much can be readily accepted, since the degree to which rationality is or is not attained by a collective is not our primary concern. Our concern is with the process by which collective rationality is attained and, more specifically, with whom it is that institutes and handles that process. The suspicion we have been pursuing is that the process by which rationality is collectivised is a process which pre-requires a group agent, and so cannot be the process through which group agents emerge. This means that the question we need to ask about the sort of collective reasoning implicit in straw vote-type procedures is whether it is best seen as an achievement of the individuals in the collective or of an already-formed group agent.

Pettit's discussion of the reason/rationality distinction as it pertains to individuals suggests that when that distinction is transposed to the social level, it must be the group itself that reasons:

> 'The rationality of the simple creature is realised sub-personally so far as there is nothing *the creature* can do in order to improve its rational performance… We reasoning creatures transcend this limitation. We do not have to rely on the processing for which our nature programs in order to be rational. *We* can do something about it… The transcendence that reasoning achieves gives [individuals] a degree of personal control over [their] own rational performance'[40]

If the exercise of reason in individuals is a personal-level phenomenon it would seem that the exercise of reason in collectives must be a group-level phenomenon – and the control over its rational performance which such exercise strives for is something the group itself achieves. Though individuals participate in the straw vote procedure through which group reasoning is enacted, their participation is not agential in the sense required for this reasoning to be a plausible joint action; they are not the primary units of agency which direct this activity and which the activity itself serves. It is the rationality of the group itself which the activity seeks to reinforce and it is the agency of group itself which does the reinforcing. The contributions of individuals to this activity are not just subvenient but *subservient* to the rational agency of the group when it is instituted and sustained by means of reasoning. This means that when Pettit claims that group agents are not born but made through a process of group reasoning he introduces a puzzle about the genesis of group agents that he does not appear to have the resources to resolve.

---

[39]List & Pettit (2011: 63)

[40]Pettit (2007: 502, italics added)

3   Conclusion: Recognition and Group Agency

Pettit's account of the coming into being of group agents depicts this process as having two distinct stages. In the first stage we find that certain collectives, whose members share a common purpose, will eventually encounter situations of the sort exemplified by the Doctrinal Paradox – situations which pressurise them into taking steps to satisfy the demands of rationality at the collective level. This pressure can be seen as the realisation on the part of members of the collective that unless they make their collectively-held attitudes rational, they will find themselves unable to effectively pursue the purpose they all share. In the second stage, the members respond to this pressure by jointly acting to commit themselves to organisational mechanisms guaranteeing the rationality of their collective views; in this way, they make themselves into an autonomous group agent. I have sought to cast doubt especially upon the coherence of this second stage in the process: it does not seem plausible to me that individuals could jointly act so as to become, effectively, subpersonal machinery in a larger person.

If I am right then Pettit's account of the emergence of group agents is not compelling, or at least is not complete. In this final, concluding section I look beyond Pettit's own theory to see whether some assistance can be brought in from elsewhere; my tentative suggestion is that this can be found in what Pettit would likely consider the most unwelcome of places.

3.1   A Hegelian Dilemma

The question we have been pursuing – 'Who makes a group agent?', or 'Who is to collectivise reason?' – is strikingly reminiscent of a question Hegel asks in the *Philosophy of Right*. Reflecting on the potential unity of 'a people', Hegel asks 'Who is to frame a constitution?'[41] The question seems to pose the following dilemma:

> 'either the would-be framers are an atomistic collection of individuals, who thus lack the moral unity that a political constitution presupposes, or they already identify themselves as a people, a *Volk*, in which case they already exist as a constituted "moral unity," whose unwritten constitution exists in its *Volksgeist*'.[42]

The first option makes the framing of a constitution impossible, while the second makes it unnecessary. Our question, 'Who is to collectivise reason?' can be seen to create a similar dilemma: either the would-be collectivisers are an atomistic heap of individuals who lack the unity of purpose that the very process of collectivising reason requires (which makes the process impossible), or they are already a unified group that can think and act as an integrated whole (which makes the process unnecessary).[43] The various efforts made by Pettit and his collaborators to finesse the being-becoming circularity – especially the suggestion that joint action can account for the emergence of group agents – can thus be read as attempts to show up this dilemma as false. But, as we have seen, the jury (or the panel of judges, as the case may be) is still out on the success of these attempts.

How does Hegel resolve the dilemma posed by his question? The dilemma is false, he claims, since it presupposes that a constitution is something which can be *made*:

> 'it is absolutely essential that the constitution should not be regarded as something made, even though it has come into being in time. It must be treated rather as

---

[41] Hegel (1967: S273)

[42] Rehg (2011)

[43] It is worth emphasizing that though they bear striking *structural* resemblances to one another, Hegel's dilemma and Pettit's should not be thought of as relating to the very same issue. Hegel is considering a very different sort of unity (moral unity, as opposed to unity of purpose and activity) and a very different sort of social creature (a 'people', as opposed to an organised, purposive collective) from that with which Pettit is concerned.

something simply existent in and by itself, as divine therefore, and constant, and so as exalted above the sphere of things that are made'.[44]

In short, the dilemma vanishes the moment we accept that a people is not something which comes to be 'constituted' by their own or anyone else's efforts; constitutions cannot be made but can only be made explicit or, perhaps, amended. It is senseless to try to institute a constitution since the only possible subjects of a constitution are those who are already constituted as a people ('sharing a Volksgeist'), and hence already the holders of a constitution (though it may be implicit). But if constitutions cannot be instituted by human effort then where do they come from? Hegel's suggestion is that they are 'divine' – that they are the product of what he elsewhere calls 'objective spirit'. What this means is that a people cannot bootstrap themselves into existence by instituting their own moral unity:

> 'we cannot be exercising an *instituting power* without taking ourselves to have been *instituted* as a "we." We cannot just institute ourselves into our instituting powers. Therefore, one has to acknowledge a primacy of... objective spirit over the subjective spirit'.[45]

3.2   A Hegelian Solution for Pettit

Is a structurally similar move available to Pettit? If he were to give up on the claim that groups are 'made, not born' and acquiesce to some analogue of the Hegelian account of the constitution of a people, would a route out *his* dilemma be forged? It certainly would, since it would allow him to say that the collectivisation of reason is something which simply happens – perhaps organically, perhaps through some or other 'divine' means – but not which something anybody (the group, or the members of a collective) has to *achieve*. And once the group is up and running ('born'), the sorts of mechanisms we have been discussing can then intelligibly feature as means by which the group sustains itself.

But what is equally certain is that Pettit would emphatically reject this sort of solution. Indeed (as noted in the introduction), one of what he considers to be the chief virtues of his account of group agency is that it can be sharply distinguished from the mysterious 'animation theory' about groups:

> 'the broadly Hegelian thought... [that] individual members give rise to a single agent only when a counterpart force comes on stream with its transforming effect'[46]

I do not wish to consider the prospects for foisting 'animation theory' upon Pettit. Instead I want to consider the extent to which a different 'broadly Hegelian' approach might be of some use to Pettit in bypassing his being-becoming trouble. The approach is inspired by Robert Brandom's[47] social pragmatist reading of Hegel's master concept of 'mutual recognition', and what I think it helps to show is how the emergence of group agents might be a more *social* process than Pettit allows. Specifically, it seems plausible to think that the recognitive attitudes of others in a wider discursive might have some role to play in the constitution of a group intentional subject.

We can begin with something which Pettit clearly accepts about the notion of a rational agent (or indeed a person): that it is a normative status. What this means is that being a rational agent (or person) implies that one is normatively constrained in certain ways and also that one

---

[44]Hegel (1967, S273)

[45]Descombes (2011: 388)

[46]List & Pettit (2011: 73-74)

[47]Brandom (2002, 2009)

is normatively entitled to adopt certain attitudes and take certain courses of action. But whom or what institutes these normative statuses? What makes these obligations and entitlements *genuine* or binding? According to Brandom, it is part of Kant's notion of the enlightenment ideal of 'autonomy' that we realise that no such normative status would be genuine were it to be imposed from without.[48]

The point can be illustrated by analogy with the chief development of enlightenment political philosophy – the idea of the social contract. Just as social contract theorists claimed that the legitimacy of political authority over citizens derives from those citizens' autonomous choice to *subject themselves* to that authority, so too do the norms which govern our thought and talk more generally only get a grip on us if they are somehow self-imposed. But this creates a puzzle, since it is hard to see how the norms of reason to which rational creatures are subject could be both self-imposed and genuinely binding: if we are the masters of our own boundedness then are we really *bound*? It is this puzzle – generated by Kant's dual insistence on normativity and autonomy – which Brandom sees Hegel as seeking to address via his notion of mutual recognition.

The solution wrought by Hegel involves the notion of a community:

> 'Hegel's principle innovation is his idea that in order to follow through on Kant's fundamental insight into the essentially *normative* character of mind, meaning, and rationality, we need to recognize that normative statuses such as authority and responsibility are at base *social* statuses'.[49]

The enlightenment ideal of autonomy entails that normative statuses are instituted by human attitudes, and for Brandom's Hegel, the key attitude is that of 'recognition' – the *treatment* of others as bearers of those statuses. What we need in order to actually be authoritative, or responsible, is to be recognised as such by others. But in seeking to be so recognised we are also recognising – we are according a certain kind of authority to those others we seek to be recognised by, namely, the authority to recognise us. According to Brandom's Hegel, normative statuses are only instituted when recognition is in this way symmetrical, or 'reciprocal'.

Reciprocal recognition thus provides not only a distinctively social account of what it is to *be* the bearer of certain normative statuses, it also provides a distinctively social account of what *becoming* the bearer of such statuses involves. Coming to be a good chess player, to borrow Brandom's own example, is a matter of coming to be recognised as such by those one recognises as such. Reciprocal recognition refers not just to an inter-subjective attitudinal *pattern* but to a certain sort of inter-subjective *process*:

> 'Achieving the status of being a good chess player is not something I can simply do by coming subjectively to adopt a certain attitude towards myself… It is up to me whom I recognise as good chess players, in the sense in which I aspire to be one. But it is not then in the same way up to me whether I qualify as one of them… To earn their recognition in turn I must be able to play up to their standards… My recognitive attitudes can define a virtual community, but only the reciprocal recognition by those I recognise can make me actually a member of it, accord me the status for which I have implicitly petitioned by recognising them'[50]

Could something like this model of reciprocal recognition in principle be exploited by Pettit to overcome the problem of how group agents emerge? I think it could. For one thing it

---

[48] Brandom (2009: Ch 2)
[49] Ibid., p. 66
[50] Ibid., pp. 70-71

need not involve any of the metaphysical mysteriousness Pettit associates with Hegelianism: for Brandom's Hegel, *Geist* or 'spirit' is nothing other than the 'realm of discursive activity' which is 'synthesized by reciprocal recognition'.[51] Perhaps then we can see reciprocal recognition as achieved in part by the joint action of members in a collective: group members can jointly act to petition a 'virtual community' for recognition as an intentional subject. If they succeed – if the community defined by their joint act of petitioning reciprocates by recognising them as an intentional subject – then their status as such is secured. Pettit's individualistic, 'performative' approach to personhood[52] obscures this possibility, making it seem as if intentional subjects could only be self-made or born, and this is what lands him with the bootstrapping difficulty I think he struggles to overcome. Group agents cannot bootstrap themselves into existence but they can take themselves to be a group agent and strive for this status to be recognised by a community which their own attitudes have delineated. If there is something to be said for this proposal then perhaps we can conclude that group agents are neither born nor made, though it may take a village to raise one.

References

Brandom, R. (2002). *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality*. Cambridge: Harvard University Press.

Brandom, R. (2009). *Reason in Philosophy: Animating Ideas*. Cambridge: Harvard University Press.

Bratman, M. (1999). *Faces of Intention*. Cambridge: Cambridge University Press.

Descombes, V. (2011). 'The Problem of Collective Identity: The Instituting We and the Instituted We' in Laitinen, A & Ikaheimo, H. (eds). *Recognition and Social Ontology*. Leiden: Brill.

Garland, D. (2001). *The Culture of Control*. Chicago: University of Chicago Press.

Gilbert, M. (1989). *On Social Facts*. London: Routledge.

Hegel, G.W.F. (1967). *Philosophy of Right*, trans. T. M. Knox, New York, Oxford University Press

Kornhauser, L. A. & L. G. Sager. (1993). 'The one and the many: adjudication in collegial courts'. California Law Review, 81:1–59

List, C & Pettit, P. (2002). 'Aggregating Sets of Judgements: An Impossibility Result,' *Economics and Philosophy* 18: 89–110

List, C & Pettit, P. (2006). 'Group Agency and Supervenience', in *Southern Journal of Philosophy* (Spindel Conference 2005), 44: 85–105

List, C & Pettit, P. (2011). *Group Agency*. Oxford: Oxford University Press.

McGeer, V. (2002). 'Developing Trust', *Philosophical Explorations* 5:1 (2002) 21-38

O'Madagain, C. (2012). 'Group Agents: Persons, Mobs or Zombies', *International Journal of Philosophical Studies* 20: 2. pp 271-287.

Pettit, P. (1995). 'The Cunning of Trust,' *Philosophy and Public Affairs*, 24: 202–225

Pettit, P. (2002). 'Collective Persons and Powers', *Legal Theory,* 8 (2002), 443–470

Pettit, P. (2003). 'Groups with Minds of Their Own', *Socializing Metaphysics: the Nature of Social Reality*, F. Schmitt, ed., Lanham, MD: Rowman & Littlefield: 167–93.

Pettit, P. (2007). 'Rationality, Reasoning and Group Agency,' *Dialectica*, Vol 61, 2007, 495-519

---

[51]Brandom (2009: 72)
[52]List & Pettit (2011: 171-2)

Pettit, P. (2009). 'The Reality of Group Agents', in *Philosophy of the social sciences : philosophical theory and scientific practice*, C. Mantzavinos, ed., Cambridge, UK; New York : Cambridge University Press: 67–91.

Pettit, P and Schweikard, D. (2006). 'Joint Actions and Group Agents', *Philosophy of Social Sciences*, 36: 18–39

Quinton, Anthony. (1976). 'Social Objects', *Proceedings of the Aristotelian Society* 75: 1–27

Rehg, W. (2011). Review of 'Recognition and Social Ontology'. *Notre Dame Philosophical Reviews*. Online: http://ndpr.nd.edu/news/27905-recognition-and-social-ontology/