Christian List, **Philip Pettit**

Episteme symposium on group agency: replies to Gaus, Cariani, Sylvan, and Briggs

Article (Accepted version) (Refereed)

Episteme Symposium on *Group Agency*
Replies to Gaus, Cariani, Sylvan, and Briggs[1]

Christian List and Philip Pettit

Episteme 9(3), September 2012

Reply to Gaus

The central theme of Gerald Gaus's commentary is the understanding of rationality underlying our book. His discussion is framed around the contrast between classical, universalist understandings of rationality, and their more recent, psychologically and evolutionarily informed rivals. He calls the former "constructivist" and the latter "ecological". He argues that our book is firmly grounded in a classical, constructivist understanding, and that this, in turn, makes it vulnerable to a number of criticisms. In particular, he argues that, despite our acknowledgment that there is no "one size fits all" organizational design for a group agent, our theory gives insufficient attention to the context-specificity of the rationality requirements a group agent should satisfy – something he suggests would be better captured by an ecological approach.

In this reply, we first explain how we see the distinction between "classical" and "ecological" understandings of rationality and relate the distinction to our project. We then address two of Gaus's more specific criticisms: his criticism of the claim that group agents must satisfy "robust group rationality" (which, technically, subsumes "universal domain" and "collective rationality"), and his criticism that our judgment-aggregation analysis relies on a problematic assumption, which he calls "proposition pluralism/structure monism". We conclude with some brief remarks about the theme of group emotions, which Gaus raises in his final section.

We should clarify one point at the outset. Gaus, like some other commentators, attributes to us a "general though qualified support for premise-based decision-making as the preferred way to cope with the problems of judgment aggregation". Although we may have become associated with a premise-based approach because of the attention we have given it in earlier work, it was never intended to be the only, or even the default, method for arriving at rational collective judgments. It is merely a particularly simple – but also rather inflexible – instance of an aggregation function that illustrates some of our theory's central points, especially the failure of propositionwise supervenience between individual and group attitudes. As we discuss in the book, real-world group agents should be expected to employ more complex and less mechanical methods of aggregation (e.g., GA, 60-62).

With this caveat in place, let us turn to Gaus's central theme, the contrast between classical and ecological understandings of rationality. For present purposes, the classical understanding is the one underlying standard normative decision theory (on the practical side) and Bayesian epistemology (on the theoretical one) – or perhaps more sophisticated

[1] We are grateful to Alvin Goldman for organizing this symposium and for helpful discussions, and to Rachael Briggs, Fabrizio Cariani, Gerald Gaus, and Kurt Sylvan for their thoughtful and generous comments on our book. In references to our book, the abbreviation GA stands for *Group Agency*.

versions of those theories. The ecological understanding of rationality has recently been championed by psychologists such as Gerd Gigerenzer. Gigerenzer and others (e.g., 2000) have argued that many behaviours and human decision heuristics that at first sight seem to violate classical conditions of rationality are in fact highly adaptive responses to certain environmental conditions. A nice example is given by the "recognition heuristic". When asked which of a given pair of cities is larger – for instance, San Diego or San Antonio – many people simply name the city they recognize. At first sight, this sounds irrational, since recognition has little to do with size. But Goldstein and Gigerenzer (2002) point out that in many human environments the frequencies with which certain cities are mentioned – say, in the media or in conversations – are very good proxies for their sizes. And indeed, they report the striking finding that, "[d]espite a considerable lack of knowledge, 100% of the Germans [in their study] answered the question [about the relative sizes of San Diego and San Antonio] correctly". Americans did less well under the same heuristic, since more of them recognized both cities. This illustrates that, although the recognition heuristic is not unconditionally rational in the classical sense, it is rational in certain environments, namely those to which it is adapted. This is what is meant by calling it "ecologically rational". Gigerenzer and others take these insights to challenge the classical idea of rationality itself.

We think, however, that this interpretation is a mistake. The notion of ecological rationality can be easily reinterpreted as a version of classical rationality, *relativized to a particular context or set of environmental constraints*. Once we specify the target environment in which the recognition heuristic is to be employed, including relevant informational and computational constraints, this heuristic may well be the most classically rational procedure for comparing city sizes. Classical statistics, for example, might tell us that, given a limited data set containing little information about city sizes in a foreign country, the frequency with which we have heard of the relevant cities is the best predictor of their sizes. In this sense, ecological rationality is simply classical rationality, *conditional on a particular environment and set of constraints*.

What is the lesson of all this for *Group Agency*? We think that the criteria by which to assess the performance of a group agent – or indeed any agent, whether a human, a dog, or a robot – and which underlie our disposition to take an intentional stance towards it must ultimately be classical. It may just so happen that whether a particular organizational structure or agential constitution ensures the fulfillment of those criteria depends on the environment or context in question. Some organizational structures may facilitate a group's rational performance in some environments but not in others, just as some robotic designs may lead a robot to behave rationally in some environments but lead to breakdowns in others. We touch on some of these issues in our discussion of feasible limits and favourable conditions in Chapter 1 of the book (especially GA, 20-21), though perhaps not in enough detail. The bottom line is that a group's organizational structure may be adapted to some environments but not to others, and it may then be said to be ecologically rational in the former environments but not in the latter. What matters for the achievement of group agency, however, is still whether, *in the relevant environment*, the group behaves broadly rationally. It may just be that different organizational

structures are best suited for achieving this outcome in different environments, which is in line with our observation that there is no "one size fits all" organizational structure.

We can now address some of Gaus's more specific criticisms. The first is that the requirement of "robust group rationality", which plays an important role in our analysis, is too strong. In our technical discussion, we do indeed interpret this requirement quite demandingly, as the conjunction of a "universal domain" condition and a "collective rationality" condition. The former requires the group's aggregation function to admit as input any possible profile of individual attitudes towards the propositions on the agenda (subject only to individual rationality constraints). The latter requires it to produce as output consistent and complete group attitudes on those propositions. Both requirements can be challenged, and relaxations of each requirement have been studied extensively in the literature on judgment aggregation, including in some of our own previous work (for a survey of that literature, see List forthcoming). This is, however, an aspect on which we may not have put enough emphasis in the book.

Universal domain can be interpreted as a requirement of "robustness to pluralism" (List 2011). It demands that the group's organizational structure be able to cope with any level of diversity among its members' attitudes, subject at most to the constraints of individual rationality. Now, in some groups or environments we may have reasons to expect individual attitudes to be more cohesive or homogeneous, for instance, when member preferences are likely to be "single-peaked", or member judgments "unidimensionally aligned". We can then get away with requiring less "robustness to pluralism" of the group's organizational structure, and still expect a rational collective performance. In other cases, because of a greater level of diversity among the members' preferences and judgments, a stronger robustness requirement may be needed.

A good analogue is the case of building regulations in different places. In California, which is an earthquake-prone region, buildings have to obey stronger robustness requirements than in London, although the underlying desideratum – stability of buildings in actual and relevant counterfactual circumstances – is the same. The only difference lies in our empirical premises about the possibility and probability of various circumstances, such as earthquakes in the case of buildings and differences in opinion in the case of aggregation.

What this shows is that while we would not want to give up the "robustness" idea underlying the "robust group rationality" requirement, there may be a case for adjusting its strength to different environments and contexts, so as to demand an unrestricted domain of admissible inputs to the group's aggregation function in some cases while admitting domain restrictions in others (for a related discussion of empirical circumstances and institutional design, see List 2007).

Similar remarks can be made about the completeness requirement that we have taken "robust group rationality" to entail. In the book, we have suggested that complete attitude formation, at least on those propositions that have been placed on the agenda for adjudication, is a necessary condition for a group's agential performance (completeness

on propositions outside the agenda is obviously not required). But here, too, we recognize some room for relaxation. As also noted in our reply to Cariani, the theory of judgment aggregation, on which we draw, provides all the relevant resources for modeling the case of incomplete attitude formation. So while we defend the central idea that "robust group rationality" is a condition for group agency, we accept that the level of robustness that is required, and the strength of the relevant rationality requirements (e.g., whether they should include both consistency and completeness, or just some weaker demands) may be adjusted from case to case.

A critic might ask whether the present remarks weaken or even undermine our case against the general feasibility of propositionwise supervenience relations between individual and group attitudes. Although the relaxation of "robust group rationality" may complicate some of our formal arguments, a careful review of the theory of judgment aggregation reveals that, even when the conditions leading to the most sweeping impossibility results are lifted, the possibilities of aggregation consistent with propositionwise supervenience (technically, "independence") are extremely limited and exclude some of the most natural methods of aggregation (from various "sequential" and "distance-based" procedures in the case of judgment aggregation, to familiar procedures such as the Borda count or the Kemeny method in the case of preference aggregation – let alone various methods involving reflective equilibrium). We conclude, therefore, that our central findings stand.

Gaus's second specific criticism is that our judgment-aggregation analysis "is only of relevance if a group regularly confronts decisions where [the individuals] disagree about the truth of the relevant propositions but the relations among the propositions have an agreed-upon logical structure"; he calls this the "proposition pluralism/structure monism" assumption. Here our response is that no such assumption – or at most a very minimal assumption – is needed.

Gaus's argument seems to be driven by his focus on the premise-based procedure, which he, in turn, applies to one of the original examples of a "discursive dilemma". In this example, all individuals agree that two premises, say 'p' and 'q', are jointly necessary and sufficient for a conclusion 'r'. So everyone accepts the background constraint 'r if and only if (p and q)'. Individuals disagree, at most, on the truth-values of 'p', 'q', and 'r'. This is a clear instance of "proposition pluralism/structure monism".

But one key respect in which the theory of judgment aggregation sketched in List and Pettit (2002) and employed in the book moved beyond previous analyses of such examples was to formulate a model in which individuals can have divergent views on *any* set of propositions, including both "atomic" propositions such as 'p', 'q', and 'r' and "compound" propositions such as 'r if and only if (p and q)' or indeed any proposition about the relationship between other propositions. The only structure that needs to be held fixed is that of the underlying logic, but this can be extremely permissive. As Dietrich (2007) has shown, the by-now standard analysis of judgment aggregation can be given for *any* logic satisfying three minimal conditions ("self-entailment", "monotonicity", and "completability"). So the only residual "structure monism" that is

required for our analysis is the one tied to the acceptance of the underlying logic, which should hardly be surprising. People can disagree about almost anything while sharing a common language and logic in which their divergent views can be expressed. It is only the existence of the latter that we need to assume for our analysis.

In conclusion, we would like to comment on a final theme raised by Gaus – that of group emotions. This is something on which we have said little in our book, though in Chapter 9 we talk somewhat metaphorically about the "character" and "soul" of a group. We agree with Gaus that the nature of group emotions is an interesting subject for further investigation, and all we would like to mention here is what our preferred methodology for this investigation would be. In line with our general approach, we think that group emotions are best analyzed in a functionalist manner: as states of the group agent – in analogy with the case of an individual agent – that play a certain functional role. Of course, different accounts of what that functional role is are possible.

Gaus seems to assume, however, that any group emotion must be a fairly straightforward function of member emotions. In his example of a committee decision on the wrongfulness of bribery in business he concludes that "[t]he group will have made the normative judgment [that certain business operations involve wrongdoing], but in ways approaching the young child and the psychopath, [where] the relevant emotional response is absent; [a majority of committee members] do not experience any such emotion, as they do not believe both premises."

Although we share Gaus's concern that some group agents – for instance, some commercial corporations – may exhibit certain behavioural patterns usually associated with psychopathy (see also Bakan 2004), we think that, from a theoretical perspective, the conclusion that group emotions must always be a straightforward function of member emotions is a little too quick. Just as the supervenience relationship between the group's preferences and judgments and the contributions of its members can be complex, so there is no *a-priori* reason why a group's emotions should be a simple function of individual emotions. And even if they were, this would require further argument and could not simply be assumed. Group agency, as we hope to have shown, is a complex social phenomenon that cannot easily be reduced to the individual level, and group emotions may well share that complexity.

Reply to Cariani

Fabrizio Cariani focuses on the epistemological aspects of the book and particularly the material in Chapter 4. He is sympathetic to the broad framework developed in that chapter and towards the underlying epistemological research programme, which proceeds by acknowledging the possibility of collective doxastic agents and investigating their properties. But he thinks that some of our more specific epistemological claims require reorientation or further development. Cariani begins by offering his own reconstruction of some central points from Chapter 4 and then raises six objections: First, our declared aim to arrive merely at possibility results about a group's epistemic capacities is too weak. Secondly, the ways in which we conceptualize truth-tracking and truth-indicating

in a group agent do not offer sufficiently clear measures for comparing the epistemic properties of different aggregation functions. Thirdly, our analysis is unduly biased towards aggregation functions that disallow incomplete collective judgments. Fourthly, the notion of a group agent's indicating reliability on a given proposition does not play the conceptual role we want it to play. Fifthly, the framework for generalizing Condorcet's jury theorem to complex, multi-proposition aggregation problems incorrectly presupposes that "the probability that an agent (individual or collective) will get all of the salient propositions right is simply determined by the reliabilities on each premise." And finally, the weighted majoritarian approach that we mention in relation to decision problems with heterogeneous individual competence levels is unsatisfactory.

We appreciate Cariani's thoughtful discussion of the epistemological programme proposed in our book, though we find that some of his objections attribute to the book – or extrapolate from it – somewhat stronger claims than we intended to make. In this reply, we briefly comment on each of Cariani's objections, and indicate where we agree with his call for further development, and where we think he may have been misled by taking our claims to be stronger than intended.

Let us begin with Cariani's first objection. Our aim in Chapter 4 is to formulate an epistemic desideratum that a group agent – indeed, any agent – must meet in order to function well in its environment, and to investigate how, if it all, a group agent can meet this desideratum. Our goal is to arrive at possibility results: we identify some salient organizational design strategies, which we call "democratization", "decomposition", and "decentralization", that may, in principle, strengthen a group agent's performance with respect to the given desideratum. The examples of concrete organizational designs that we discuss, however, are purely illustrative and not exhaustive. As we point out:

> "Whether or not each of these design principles [i.e., democratization, decomposition, and decentralization] is helpful, and to what extent, depends on the group agent and epistemic task in question, and there may not be a 'one size fits all' organizational design that is best for all group agents and all epistemic tasks. But the mere possibility of the three types of benefits reinforces the potential of epistemic gains through group agency" (GA, 102-103).

We still think this passage captures our aim accurately. It may be a modest aim, but it is still an important one. The idea that some groups can constitute doxastic agents in their own right that are capable of tracking, as well as indicating, the truth in the beliefs they collectively form is not yet the established orthodoxy, and hence any possibility results of the kind we present matter.

Cariani says that "if the possibility result were the only goal of the chapter, I would happily declare myself convinced of it by a relatively small subset of L&P's arguments", and goes on to say that "for [some of the more detailed] discussion to be relevant, [L&P] must be aiming for something more ambitious". He then suggests that, over and above defending a group agent's possibility of meeting the epistemic desideratum, we might also be claiming, firstly, that "the cases of which the assumptions [i.e., those discussed in

Chapter 4] hold true are in some sense especially significant", and secondly, that "as we increase the class of circumstances that we consider, the results are sufficiently robust to give us optimism about the epistemic value of pooling information in groups". He is sympathetic to the second claim but skeptical of the first.

We are, however, not committed to either of these claims in unqualified form, and to the first one even less than to the second. As we emphasize in the book, real-world group agents and the decision problems they face are much more complex than any simple formal model can capture, and the usefulness of formal models and results of the kind presented in Chapter 4 lies merely in drawing our attention to certain possibilities and patterns in particularly distilled form. For this reason, the aggregation functions and hypothetical group agents we discuss to illustrate the organizational design strategies of "democratization", "decomposition", and "decentralization" are simplified ideal types, and real-world applications of the same organizational design strategies will no doubt take more complex forms, whose empirical details require further investigation.

Cariani's second objection is that our conceptualization of truth-tracking and truth-indication in a group agent does not provide sufficiently clear measures for comparing the epistemic properties of different aggregation functions. He says: "It is not clear how the reliability analysis should ground comparisons among rules". In response, we would like to make both a concession and a clarification. The definitions of truth-tracking and truth-indication given in Chapter 4 are indeed insufficient by themselves to yield a fully fledged ranking of different aggregation functions in an order of epistemic strength. But there are two important things to note here.

First, defining the *concepts* of truth-tracking and truth-indication is not the same as devising a *practical measure* of each; the latter task goes well beyond the former. To give an analogy, it is one thing to define the *concept* of knowledge (an epistemological task); it is another to devise a *practical measure* of how knowledgeable a given agent is (a more psychological task). Our focus in Chapter 4 has been on the former, conceptual exercise, rather than the latter, measurement one. At best, the definitions of truth-tracking and truth-indication that we give suggest some *heuristics* for comparing different organizational designs, and it is only in this weaker spirit that we employ them. Secondly, the units of epistemic comparison must be fully constituted group agents or organizational structures *relative to a specific set of individuals with a specific individual reliability profile*. Aggregation functions taken in isolation cannot be assessed epistemically, since their epistemic performance depends crucially on the reliability and epistemic performance of the individuals whose judgments are being aggregated. The same aggregation function can perform well in one group, and badly in another.

Cariani discusses some proposals on how the epistemic performance of different aggregation functions could be compared, relative to a given set of assumptions about how reliable the individuals are. We welcome Cariani's call for further work on this measurement problem, and wish to emphasize just two points. First, the ranking of different organizational designs in an order of epistemic strength may depend not just on their reliability properties, but also on some other objectives, for instance the costs of

different types of error. As noted in Chapter 4, when false positives and false negatives are equally costly, majority rule may epistemically outperform unanimity or supermajority rules, but when there is an asymmetry between these two types of error – say, false positives are more costly than false negatives – this ranking may be reversed. Secondly, it is often difficult to capture the epistemic performance of a given organizational design in terms of a single summary figure, such as Cariani's measure of "global competence". Since this measure is an average of a group's reliability across different circumstances, with each circumstance weighted by its probability of occurrence, the measure is highly sensitive to the assumed base rate of different circumstances. To see why this is problematic, take our example of a medical advisory panel that always certifies the safety of any chemical under investigation, regardless of the chemical's actual risks (GA, Chapter 4.1). So long as the base rate of dangerous chemicals is relatively low, the panel's "global competence" could be quite high, despite the panel's low negative truth-tracking reliability. This is the reason why, contrary to Bovens and Rabinowicz (2006), we insist on presenting judgmental reliability information on a given proposition always as a *pair* consisting of both a positive and a negative reliability, rather than just as a single average figure.

Cariani's third objection is that our analysis is unduly biased towards complete aggregation functions, i.e., aggregation functions that disallow indecision on any proposition on the agenda. We may indeed have misled readers by putting a lot of emphasis on the case of complete collective judgments, due to our focus on agents who are constrained to form attitudes on at least those propositions that are placed on the agenda (evidently, incompleteness on propositions not on the agenda is no problem). However, the theory of judgment aggregation, on which the book draws, has explored the case of incomplete collective judgments in detail, beginning with our own initial paper (List and Pettit 2002, Section 4) and subsequently in Gärdenfors (2006), Dietrich and List (2007, 2008, forthcoming), and Dokow and Holzman (2010). We have just not emphasized that aspect of the theory of judgment aggregation in the book. Nonetheless, it is important to stress that our reliability analysis in Chapter 4 is not by itself biased against incomplete aggregation functions. To see this, recall that the costs of different types of error may enter our assessment of an agent's epistemic performance. Now consider the case of an expert panel whose credibility depends on avoiding the assertion of falsehoods. It may well be appropriate for such a panel to use a (symmetrical) supermajority rule whereby it collectively accepts only those propositions that command a strong supermajoritarian support among the experts. In this case, the group's epistemic performance is optimized by an aggregation function that frequently delivers incomplete collective judgments.

Cariani's fourth objection concerns the notion of indicating reliability. We define an agent's positive indicating reliability on a proposition 'p' as the conditional probability that 'p' is true given that the agent judges that p, and the negative indicating reliability as the conditional probability that 'p' is false given that the agent doesn't judge that p. We also point out that an agent's positive, or negative, indicating reliability can be naturally interpreted as the probability that an outside observer is entitled to assign to 'p' on learning that the agent has, or has not, judged that p. Cariani challenges this

interpretation, arguing that the outside observer should look not just at what the agent has, or has not, judged, but also – at least in the case of a group – at the breakdown of individual votes leading to the relevant collective judgment. We don't deny that, when the breakdown of individual votes is publicly available, these votes may carry some information, and an observer's Bayesian belief update should take that into account. Indeed, drawing on earlier work in List (2004), we state a formula in Chapter 4 that captures the conditional probability that 'p' is true, given a particular pattern of votes for and against 'p', under Condorcetian assumptions.

However, just as the neural processes leading to an individual's belief formation are usually opaque to an outside observer, so the individual-level history of a group agent's judgments may often be opaque to the outside world, either in practice or even by design. We take an agential stance towards the collective system as a whole, and when we assess the truth-tracking and truth-indicating abilities it possesses as an integrated unit, the definitions of truth-tracking and truth-indication, initially defined for an individual agent, apply naturally to the collective as well. Even in cases in which the group members' voting pattern is public, we may still be interested in the indicating reliability of the group agent as a whole. For example, we have good normative reasons to be interested in the indicating reliability of the Supreme Court as a whole, and not just in the information carried by the individual judges' votes.

Cariani's fifth objection is that "[t]he framework for generalizing the Condorcet-style analysis to complex aggregation problems assumes that the probability that an agent (individual or collective) will get all of the salient propositions right is simply determined by the reliabilities on each premise" – let us call this the "compositionality assumption" – and this "seems to [him] to be relatively uncommon". We agree with Cariani that an agent's probability of making correct judgments on all relevant propositions is not always a function of the agent's positive and negative reliabilities on some relevant premises. Even fairly simple factors such as stochastic dependencies between someone's judgments on different premises could undermine the compositionality assumption. Furthermore, the agenda of propositions under consideration need not be neatly divisible into a set of epistemically prior premises and a set of epistemically subordinate conclusions. Nothing in our approach to formulating the epistemic desideratum, or in our suggestion that democratization, decomposition, and decentralization are ideal-typical strategies for improving a group's epistemic performance, hinges on the compositionality assumption. We are not committed to that assumption and agree with Cariani that it is not generally true. He is right, of course, that, without compositionality, a Condorcetian analysis of a group's truth-tracking ability in multi-proposition cases becomes more complicated, and there is certainly scope for further work here.

Cariani's final objection concerns the remarks we make about weighted majority voting in the case of heterogenous individual reliability. We mention in Chapter 4 that even when one of the jury theorem's central assumptions – namely that all individuals are equally competent on a given proposition – is violated, a weighted form of majority voting can still epistemically outperform a dictatorship or other exclusionary decision method. To make this point, we invoke the statistical result that weighted majority voting

maximizes the group's positive and negative tracking reliability if each individual's vote is given a weight proportional to the log-likelihood ratio associated with the individual's reliability (e.g., Grofman, Owen, and Feld 1983, Ben-Yashar and Nitzan 1997). Cariani finds this an unsatisfactory "solution" to the problem of heterogenous competence. But it was never meant to be a practical solution. Again, we presented the result as a possibility result: to show that even the relaxation of some of the idealized assumptions of Condorcet's jury theorem does not rule out the in-principle availability of the epistemically advantageous organizational design strategies identified in the chapter, here the strategy of democratization. We completely agree with Cariani, however, that the *practical* question of how to make collective judgments in cases of heterogenous competence – especially when the competence levels of different individuals are opaque or even normatively deemed irrelevant – is a difficult question that requires further research.

Reply to Sylvan

Kurt Sylvan goes along with the realism about group agents that we defend but challenges our claim that this realism turns group agents, and their attitudes, into non-redundant posits of common sense, social science, and normative theory. In responding to his challenge, we look first at what it means to claim non-redundancy for these posits and then consider Sylvan's argument. While we end up holding our ground, we appreciate the opportunity to consider his interesting and novel challenge; it raises some great issues.

The non-redundancy claim in *Group Agency* involves two points. First, the attitudes we ascribe to a group agent are "not readily reducible" to the attitudes of members (GA, 5): that is, talk about them is not "readily translatable into individualistic terms" (GA, 6). And second, the ascription of such attitudes gives us a "distinctive way of understanding and relating" to the group; it enables us "to interact with it, criticize it, and make demands on it, in a manner not possible with a non-agential system" (GA, 5-6). According to the theory developed, there is nothing in a group's counting as an agent that is not supervenient on the attitudes and acts of the members, just as, according to physicalism, there is nothing in an individual's counting as an agent that is not supervenient on its physical constitution. But still, our theory implies that the move to recognizing a group as an agent gives us a novel perspective of the kind that many physicalists associate with seeing a material system as an intentional agent. In each case, "the new perspective enables us to see higher-level regularities" and "to respond to the system … in abstraction from much of what happens at the lower level" (GA, 6).

There is no clean divide between redundant and non-redundant forms of realism about entities that are composed out of other entities, as group agents are composed out of members and individual agents out of physical states. Towards one end of the spectrum there are realist theories according to which reduction or translation from the higher to the lower level is easy and the response capacity associated with the higher stance does not strictly require that perspective. Towards the other end are realist theories according to which the reduction is harder and the response capacity supported by the higher stance is more or less exclusively tied to it. Towards the first extreme, intuitively, there must

exist a workable, if complex, recipe for determining on the basis of lower-level information when it is appropriate to make higher-level attributions; towards the second, no such recipe is in sight. Citing problems that raise mounting difficulties for reduction (GA, 77-78), we argue that our theory of group agency lies towards this latter end.

Perhaps miscued by some of our remarks, Sylvan treats the question of whether certain posits are redundant as an on-off matter rather than one of degree. And he sets the bar for non-redundancy quite high, arguing that the issue is whether there is "an uncloseable explanatory gap" between the lower and the higher levels of discourse, where that gap is of the kind physicalists generally acknowledge as a problem for making physical sense of consciousness (though not of intentionality). He invokes a comment we make on parallels with the discussion of consciousness (GA, 75) to suggest that we take the same view of the nature of the issue. But we differ with him on these matters. We think that whether a theory counts as introducing redundant or non-redundant posits is a matter of degree, and we do not think it depends on whether there is an explanatory gap of the kind discussed in the consciousness literature.

We now put aside these differences about how to construe the issue between redundant and non-redundant realism and turn to Sylvan's claim that we have not established anything that might deserve to be called a non-redundant realism. According to our theory, the beliefs of a group agent in a set of propositions need not be determined, proposition by proposition, on the basis of the individual beliefs of a majority of members in those propositions (or in the appropriateness of the group's believing them). Nor, more generally, need they be derivable by any systematic, propositionwise function from the individual beliefs of members. It is this "autonomy" of group beliefs (or attitudes in general) that lies behind our claim that group agents should be regarded as agents in their own right, with attitudinal sets that may be strikingly autonomous from the attitude sets of their members.

While accepting this core thesis, Sylvan concentrates on its illustration under an inflexible premise-based mode of aggregating judgments. Under a premise-based procedure, the members of a group treat certain mutually independent propositions as premises, collectively endorsing them if and only if they command majority support. But when it comes to a proposition, 'p', entailed by the premises that they have already endorsed, the group is prepared collectively to endorse 'p' even when it turns out not to enjoy majority support: even perhaps when everyone disbelieves it. The possibility of a logically derived but individually unsupported group belief in a proposition like 'p' illustrates (but is only one among many possible illustrations of) how a group belief in a proposition need not be a majoritarian function — nor, more generally, any other function — of the beliefs of members *in the same proposition*.

In our book we dismiss the premise-based procedure as a satisfactory way of organizing a group agent on the grounds that, first, it presupposes that members can agree on what should count as premises and, second, it does not permit a group to reconsider any premise-beliefs as a result of seeing the conclusions that would follow from them (GA, 60-62). In more plausible and workable group agents, the organizational structure will

allow any proposition, however long endorsed, to be rejected on grounds of inconsistency. And it may admit different modes of determination for different kinds of proposition, authorize different subgroups to specialize in determining the group's beliefs on different issues, and rely on any of a range of devices to ensure that the group's overall beliefs are coherent. Arrangements of this sort are familiar from everyday bodies like companies and churches, voluntary associations, trade unions, and political parties.

Acknowledging that group agents are liable to have novel attitude sets, and can count as agents in their own right, Sylvan argues that nonetheless there is a ready reduction of any group belief (or analogously, group preference) in an individually unsupported conclusion — say, 'p', in our example — despite the fact that the belief does not reflect the members' individual beliefs in 'p'. In a case like this, as he puts it at one point, the group belief that p "is constituted by a pattern of members accepting it *qua* members" — that is, by their treating it as true "for group purposes" — at least where this pattern is a matter of common knowledge. And holding that in this sense "group belief plausibly boils down to neat patterns of individualistic facts", he thinks that our realism about group agents falls on the redundant side of the spectrum.

But this argument depends on the assumption that there is a relevant parallel between the way an individually supported group belief in a proposition relates to member *beliefs* in that proposition and the way an individually unsupported group belief in a proposition relates to member *acceptance* of the proposition. The idea is that just as underlying member belief would make the first sort of group belief unsurprising, so underlying member acceptance would make the second sort of group belief equally unsurprising. But the suggested parallel is spurious and does not establish that group attitudes and agents are readily reducible to individualistic facts.

Everything a group does, it does by means of individuals, no matter what structure it has; it is a superordinate entity in which individuals are the only moving parts. Thus when a group comes to form a belief, individually supported or not, it is essential that relevant members accept it. More specifically, those who are authorized to perform in a judging role must assent in the name of the group to that proposition, and those who are authorized to perform in executive roles (they may or may not be the same individuals) must be led to act on the proposition in the name of the group. These executive members have to act in their respective roles as if the proposition is true but they need not personally assent to it; provided their action fits with the general pattern required of the group, they may be guided only by unexplained cues or instructions.[2]

This observation shows that in the role-based sense of acceptance in which it means treating a proposition as true for group purposes, the relevant members – those enacting the group attitudes – have to accept all the propositions the group believes, individually supported or not, whether in the sense of assenting to these propositions or merely acting on them. But this does not mean that the group beliefs can easily be reduced to – or

[2] With some propositions that members accept, as in the propositions corresponding to rules of inference that they never spell out, the only mode of acceptance on anyone's part will be the executive. Lewis Carroll (1895) shows that there have to be some such rules and some such propositions.

reductively explained in terms of – the pattern of role-based member acceptances of these propositions. While all of a group agent's beliefs require the relevant *acceptance* of relevant members, not all have to be a majoritarian or other function of any member *beliefs* in their contents. It is this aspect of group agency that we stress in arguing that while all the attitudes of a group agent have to be supervenient on the contributions of members — and can be supervenient on any of a variety of contributory profiles — only some can satisfy the sort of reducibility that would obtain under a majoritarian or other propositionwise function.

Were the beliefs of a group agent a simple function — say, a majoritarian function — of the members' beliefs in the relevant propositions, then the recipe for attributing group belief would be entirely clear: let a majority of members believe any proposition, and the group counts as believing it. But knowing that the beliefs of a group have to be a function of the acceptance that members give to those propositions offers us no such recipe. Put in somewhat oversimplified terms, it is the group belief that explains the role-based member acceptance, not the role-based member acceptance that explains the group belief. And so the individualistic information about acceptance does not offer a base for the ready reduction of group attitudes. To quote from Rachael Briggs's commentary, it directs us to "a complicated and disjunctive proposition about a wide range of individual attitudes".

A little reflection makes clear why this is so. In order to know that members accept a certain proposition we must know who assents to it on behalf of the group and who acts on it in the name of the group. In order to move from knowledge of such member acceptance to knowledge about a putative group belief we must know that this pattern is one that fits with the organizational structure under which members operate: it is relevant members, for example, who give relevant forms of acceptance. And in order to know that the group actually believes the proposition — that it meets the functional specifications on belief — we must know that the structure allows the group to perform as an agent, letting the putative belief serve in the role proper to belief. The individualistic knowledge base required for determining group beliefs and other attitudes is just too complex to allow anything that could count as a ready reduction.

In developing his challenge, Sylvan chides us for not paying more attention to the literature on joint intention and action on the grounds that various writers in that tradition offer more or less individualistic recipes for when to ascribe a joint intention to a collection of individuals to pursue various actions and effects together, including the effect of endorsing a proposition in common. As the book stresses, however, we think that our enterprise presupposes that there is an acceptable account of what it is to act on a joint intention and does not need to compete with the established accounts in this area; here we can simply rely on the rich existing literature. We concentrate on the distinct topic of what is needed if people are to act together (in the normal case, on the basis of a joint intention) so as to constitute an agent with the capacity for rational coherence and interpersonal answerability that human agency requires.

In particular, we are interested in how people can do this, given that coherence and answerability rules out relying on any systematic, propositionwise function from

individual to group attitudes. We think that we have established the possibility of group agency in this sense, a point on which Sylvan agrees. But we also think that our account rules out the prospect of a workable recipe for using individualistic information alone to determine when it is appropriate to ascribe this or that attitude to a group agent. The group agents that people routinely posit in common sense, in social science, and in normative theory are not only real entities; they are also posits that no plausible, accessible recipe can render redundant.

Reply to Briggs

Rachael Briggs is broadly sympathetic to the viewpoint of the book, and her piece consists mainly in some useful clarifications and some virtuoso developments of a number of ideas. She makes clearer than perhaps we did that in discussing the responsibilities and rights of group agents, we are focusing on the responsibilities and rights they ought to be given under established conventions: under social norms and laws. She identifies the sort of argument that ought to be made in favor of the normative individualism that we take for granted: this is, roughly, the thesis that in determining suitable norms and laws we ought to consult the interests of individual human beings, not group agents. She challenges our contention that such a normative individualism is needed for putting a limit on the rights of group agents. And she develops an alternative response to the objection that groups can't be in control of the things they do because members are already in control, and hence that groups cannot be held responsible. Our response in the book relies on Jackson and Pettit's model of programme explanation; Briggs develops a response based on the work of List and Menzies on difference-making causation.

In this reply to her commentary, we focus mainly on the issue of corporate rights and responsibilities, but we should first make a brief comment on an early remark to the effect that we provide only "an argument that group agents are possible, not an argument that group agents exist". The point may not be very important given, as Briggs notes, that it is "overwhelmingly likely that group agents exist to fill the conceptual space … carved out". But it may be useful to clarify what the book claims on this front.

We sketch an argument for the existence of group agents in the following passage from the end of Part I, omitting a crucial empirical premise.

> "Let a collection of individuals form and act on a single, robustly rational body of attitudes, whether by dint of a joint intention or on some other basis, and it will be an agent. The argument of the second and third chapters shows that despite the difficulties of aggregation, it is possible for collections of individuals to coordinate their individual contributions so as to achieve this level of functioning. Hence group agents exist." (GA, 75)

The missing premise is, of course, that as a matter of empirical fact there are collections of individuals who do form and act on suitable bodies of attitudes. That premise we mostly take for granted in the book, though we provide some anecdotal support for it in

our brief overview of the variety of attitudinally organized groups, political, commercial and civic, that are recognized in common sense and in social science (GA, Chapter 1).

We recognize that our book does not primarily focus on this empirical premise and that in our discussion of ontology, in the first part of the book, we look mainly at the possibility of group agents. Furthermore, when the book discusses the design and status of group agents, our interest is primarily in laying out the conceptual territory rather than in empirically mapping various real-world group agents and their capacities. Thus we entirely agree with Briggs's comment that there is a need for further empirical "investigation into the actual dynamics of group reasoning to establish which agential capacities groups have, and how group attitudes relate to the attitudes of group members".

Turning now to the issue of corporate responsibilities and rights, the first thing to note is that, as Briggs observes, we take the view that "there is every reason to hold [group agents] responsible in their own right", but that they should "have restricted rights as compared with individuals" (GA, 182). In other words, there is a difference between the way we treat groups as the bearers of responsibilities and the way we treat them as the bearers of rights. As Briggs puts it, on this account, "there is an important asymmetry between obligations and rights".

We now feel that our comments in the book on this kind of asymmetry may be somewhat misleading. In discussing responsibilities, we argue that a group agent is fit to be held responsible for what it does in its own name, but we do not pay much attention to the range of responsibilities that it ought to be given: the range of actions for which it is fit to be held responsible. In discussing rights, on the other hand, we take for granted that group agents are fit to be given rights in their own name — that is, rights over and above the rights accruing to members — and focus rather on the range of rights they should be given. We might, however, also express the view by saying that group agents are *both* fit to be held responsible *and* fit to be given rights in their own name, but that there are serious limitations on the rights they should be given — as, presumably, there are some limitations on the responsibilities they can plausibly have.

Briggs considers our arguments for why group agents are fit to be given responsibilities and rights, elaborating in a useful way on the various considerations that are relevant to this. On fitness to be held responsible, she observes that we might have drawn on the work of List and Menzies on difference-making causation in order to counter the objection that it is members, not groups, that have control of what groups do; this is a point we are aware of, but didn't develop in the book. And, adding some further observations of her own, she endorses the argument we provide for why group agents satisfy the different conditions associated with being fit to be held responsible for a given action.

Addressing a group agent's fitness to be given rights, Briggs focuses on the performative conception of persons that we adopt, according to which, in her words, "to be a person is to be capable of the understanding and competence required for participation in a social

contract". While she thinks that it may be better to adopt "an interest-based conception of personhood", she acknowledges in our defense that "[t]his knowledge and understanding make it appropriate to grant conventional rights to an agent", including a group agent. Let a group agent be capable, as we argue, of being "party to a system of accepted convention, such as a system of law" (GA, 173) — specifically, a "knowledgeable and competent" party to the system — and it is clearly capable of having rights: that is, capable of enjoying, recognizing, and even invoking the protection of certain choices that any relevant system of convention or law will provide.

We have been discussing a group agent's fitness to be held responsible and its fitness to bear rights. Turning now to the other question that arises in each case, what range of responsibilities and rights ought a group agent to be given? In our book, we do not explicitly address the range of responsibilities in question, taking for granted that it will make sense to assign responsibilities to a group agent only in the domain where it is equipped to form and act on attitudes – that is, where it meets the conditions for fitness to be held responsible. But we say in passing that the question can be resolved consistently with the same normatively individualist basis as the range-of-rights question, remarking that our views on both sides "make sense within the perspective of normative individualism" (GA, 182).

Briggs's main challenge arises at this point. Our normative individualism suggests that in determining what responsibilities and rights a group agent ought to be given, we should look only to the interests of human beings: "something is good only if it is good for individual human or, more generally, sentient beings" (GA, 182). She takes the discussion to a deeper level, however, in distinguishing the different interests of individual human beings that might be invoked; in arguing that there are counterparts to those interests at the level of group agents; in suggesting that there are grounds for taking such group interests into account as well; and finally in maintaining that consistently with doing this we may still be able to stick with the restrictions we support on the rights that ought to be given to group agents.

Briggs follows Parfit in arguing that something may be good for an agent — for short, may be in that agent's interest — in virtue of its affecting the agent's hedonic satisfaction, the fulfillment of the agent's desires and plans, or the agent's enjoyment of certain allegedly objective goods like knowledge, friendship, or freedom. She then maintains that on our realist theory, group agents as well as individual agents may be capable of having interests of these three kinds. We agree with her that group agents may have interests of the last two kinds, but we remain agnostic about the possibility of interests of the first kind. Briggs claims that "it's not so obvious that there *couldn't* be a group that felt pain". While we have certainly argued that group agents can have relatively autonomous intentional states, which are not easily reducible to the members' corresponding states, it is an open question whether group agents could also have hedonic states, and if so, whether these could ever go beyond something that is straightforwardly reducible to the corresponding states of members.

Still, given that group agents do have interests of non-hedonic kinds, the question is why these shouldn't weigh against the interests of individuals in determining what is for the best in social life: in particular, what allocation of corporate rights and responsibilities is for the best. Our statement of normative individualism suggests that the interests of group agents count for nothing, but as Briggs makes clear, this is far from obvious. They might count equally with the interests of individuals, they might count but have a weaker weight than the interests of individuals, or they might count but only in determining the best candidate among arrangements that do equally well in individual terms. Were we to go along with the interest-based conception of persons which Briggs herself finds attractive, this result would be reinforced. On that view, persons are just those creatures "for which things can go well or badly". We do not go along with this conception for the methodological reason she gives: "the interest-based conception blurs [the descriptive-normative] distinction; the idea of interests is already a normatively, perhaps morally, loaded concept". But even without an interest-based conception of personhood, it is certainly plausible that group agents can have interests, and it is not clear why these should not count at all in determining the best social arrangements.

We are happy to concede to Briggs that if the interests of group agents are allowed to count — if normative individualism is false — the danger that group agents represent for the welfare of individuals, which we ourselves invoke, would still argue for giving fewer rights to group agents than to individuals. But is it reasonable to stick with normative individualism and to hold the line against her main challenge?

This is a difficult question and requires far more investigation than we give it in the book, where we adopt normative individualism as just a plausible assumption. But there is one consideration that might count in favor of normative individualism and which we gesture at in our discussion. This is that group agents, unlike individuals, are the products of social arrangements and depend both for coming into existence and continuing in existence on the coordinated efforts of their members. "Individual persons create and organize group agents" (GA, 181), as we say in the book. There may seem to be something of a circle involved in thinking that, even when we hold the population of all individuals fixed, the criterion by which we should make a judgment on the best social arrangements involves the interests of social entities that those very arrangements bring into existence. It seems more natural to think that it is the interests of individuals that determine what sorts of rights and responsibilities group agents are to be given; individual agents are the producers of social arrangements, corporate agents the precipitates.

When individuals create and sustain group agents, whether in political, commercial, or civic spheres, then the rights and responsibilities given to those corporate bodies have important repercussions for all those affected, members and non-members alike. Some individuals may benefit and others suffer from the form such rights and responsibilities assume. Whatever ethical view we adopt, it seems plausible to think that the interests of the individuals who stand to benefit or suffer as a result of social arrangements should ultimately determine the rights and responsibilities to be assigned, not the interests of the corporate entities that these arrangements bring into existence.

References

Bakan, J. (2004). *The Corporation: The Pathological Pursuit of Profit and Power*. New York (Free Press).

Ben-Yashar, R. and S. Nitzan (1997). "The optimal decision rule for fixed-size committees in dichotomous choice situations: the general result." *International Economic Review* 38: 175-186.

Bovens, L. and W. Rabinowicz (2006). "Democratic Answers to Complex Questions: An Epistemic Perspective." *Synthese* 150(1): 131-153.

Carroll, L. (1895). "What the Tortoise said to Achilles." *Mind* 4: 278-80.

Dietrich, F. (2007) "A generalized model of judgment aggregation." *Social Choice and Welfare* 28(4): 529-565.

Dietrich, F. and C. List (2007). "Judgment aggregation by quota rules: majority voting generalized." *Journal of Theoretical Politics* 19(4): 391-424.

Dietrich, F. and C. List (2008). "Judgment aggregation without full rationality." *Social Choice and Welfare* 31: 15-39.

Dietrich, F, and C. List (forthcoming). "Judgment aggregation with consistency alone." Working paper, London School of Economic, available at: http://personal.lse.ac.uk/list/PDF-files/ConsistencyAlone.pdf

Dokow, E., and R. Holzman (2010). "Aggregation of binary evaluations with abstentions." *Journal of Economic Theory* 145: 544-561.

Gärdenfors, P. (2006). "An Arrow-like theorem for voting with logical consequences." *Economics and Philosophy* 22(2): 181-190.

Gigerenzer, G., P. M. Todd, and the ABC Research Group (2000). *Simple Heuristics That Make Us Smart.* New York (Oxford University Press).

Goldstein, D. G., and G. Gigerenzer (2002). "Models of Ecological Rationality: The Recognition Heuristic." *Psychological Review* 109(1): 75-90.

Grofman, B., G. Owen, and S. L. Feld (1983). "Thirteen theorems in search of the truth." *Theory and Decision* 15: 261-278.

List, C. (2004). "On the significance of the absolute margin." *British Journal for the Philosophy of Science* 55: 521-544.

List, C. (2007). "Deliberation and agreement." In Shawn W. Rosenberg (ed.), *Deliberation, Participation and Democracy: Can the People Govern?* Basingstoke (Palgrave).

List, C. (2011). "The Logical Space of Democracy." *Philosophy and Public Affairs* 39: 262-297.

List, C. (forthcoming). "Judgment aggregation: an introductory review." *Synthese*, in press.

List, C. and Pettit, P. (2002). "Aggregating Sets of Judgments: An Impossibility Result." *Economics and Philosophy* 18: 89-110.