

ARTICLE

Responsibility Incorporated*

Philip Pettit

The *Herald of Free Enterprise*, a ferry operating in the English Channel, sank on March 6, 1987, drowning nearly two hundred people. The official inquiry found that the company running the ferry was extremely sloppy, with poor routines of checking and management. "From top to bottom the body corporate was infected with the disease of sloppiness."¹ But the courts did not penalize anyone in what might seem to be an appropriate measure, failing to identify individuals in the company or on the ship itself who were seriously enough at fault. As one commentator put it, "The primary requirement of finding an individual who was liable . . . stood in the way of attaching any significance to the organizational sloppiness that had been found by the official inquiry."²

In a case like this it can make good sense to hold that while the individuals involved may not bear a high degree of personal responsibility, together as a corporate enterprise they should carry full responsibility for what occurred. Although the members may not fully satisfy the conditions for being held personally responsible – although there are mitigating circumstances that excuse them in some measure – the organization as a whole may satisfy such conditions perfectly and may be fully fit to be held responsible. No incorporated agency without incorporated responsibility, and

* I am grateful for comments received when the article was presented in Lisbon, at the 2005 Congress of the European Society for Analytical Philosophy, as well as at presentations in Oxford University; Cambridge University; Yale University; Australian National University; University of California, Berkeley; Harvard University; and the University of Groningen. I was led to rethink some points by anonymous referees and editors and by exchanges with John Braithwaite, Owen Fisse, Shelly Kagan, Doug MacLean, Victoria McGeer, Jeff McMahan, Wlodek Rabinowicz, Ed Rock, Gideon Rosen, Sam Scheffler, Simon Stern, Martin Van Hees, Jay Wallace, and Susan Wolf. The article is connected to a joint project with Christian List on group agency, and I have particularly profited from our many, continuing exchanges. This article was originally published in 117 *Ethics* (2007): 171–201. Republished in this journal with the kind permission of the University of Chicago.

1 Eric Colvin, "Corporate Personality and Criminal Liability," *Criminal Law Forum* 6 (1995): 17.

2 Ibid., 18.

this, even when individual responsibility is diminished. That, in a slogan, is the line I defend in this article.³

The article is in six main sections. In Section I, I give an account of what is involved in holding any agent responsible, identifying three necessary conditions for fitness to be held responsible. In Sections II–IV, I argue in turn that corporate agents often meet these conditions and so are fit to be held responsible. And then in Section V, I show why there is a point in holding corporate agents responsible, side by side with the individual agents who act in their name. The article concludes with a brief discussion of how far the lesson defended applies to looser collections of individuals, such as national and religious traditions, and not just to properly corporate agents.

What do I mean by corporate agents, or as I shall also say, corporations, organizations, and group agents? Examples include companies, parties, churches, and universities but also partnerships, voluntary associations, and town meetings. These entities operate via their members, although they retain their corporate identity through continuous changes of membership. Specifically, they operate through their members in such a way that they simulate the performance of individual agents. They endorse certain goals and methods of reviewing goals and certain judgments and methods of updating judgments, and they follow procedures that enable them to pursue those goals in a manner that makes sense according to those judgments. Sometimes they do this under the participatory arrangements that are typical of the voluntary association, but just as often they do it under the more hierarchical organization of the church or a commercial firm. We shall be looking into detail later.

I. FITNESS TO BE HELD RESPONSIBLE

Holding Responsible and Related Attitudes

There are many ambiguities surrounding the notion of holding someone responsible, and it may be useful to guard against these from the outset. The first thing to say is that holding an agent responsible for something in the sense intended here does not mean just assigning causal responsibility for what was done. We might

- 3 The point of view that I defend is shared in broad outline, although not always on specific points, with a number of recent authors, in particular with Peter French. See Peter A. French, *Collective and Corporate Responsibility* (New York: Columbia University Press, 1984); Peter A. French, Jeffrey Nesteruk, and David Risser, *Corporations in the Moral Community* (New York: Harcourt Brace Jovanovich, 1992). See too the papers in Toni Erskine, *Can Institutions Have Responsibility? Collective Moral Agency and International Relations* (London: Palgrave, 2003). But my argument is original, I think, and in any case the point of view taken remains a minority position. The modern contractualist analysis of the commercial corporation, so it appears, still treats it as little more than “a collective noun for the web of contracts that link the various participants” (Ross Grantham, “The Doctrinal Basis of the Rights of Company Shareholders,” *Cambridge Law Journal* 57 [1998]: 579). This analysis may take a number of forms, whether on the basis of identifying the corporate organization with its chief officers or treating the agent involved as a delegate or imputing vicarious or associative liability. See Peter Cane, *Responsibility in Law and Morality* (Oxford: Hart, 2002); Larry May, *The Morality of Groups: Collective Responsibility, Group-Based Harm, and Corporate Rights* (Notre Dame, IN: University of Notre Dame Press, 1987).

Philip Pettit

hold the dog responsible for soiling the carpet in the purely causal sense, but we would not hold the dog responsible in the sense that is relevant here, not at least by my understanding of canine capacities. Holding responsible in the relevant sense has as an implication that, if what was done is something bad, then the agent is a candidate for blame; if what was done is something good, then the agent is a candidate for approval and praise. We may be angry at the dog, we may even be frustrated at the failure of the training regime, but by most lights treating the dog as fit for blame will be no more sensible than treating the weather that way.

Holding someone responsible, then, is different from the mere assignment of causal responsibility. Second, it is distinct from just holding an agent accountable – identifying that person as the one who carries the can, the one who sits at the desk where the buck stops. We might hold someone accountable in that sense, and apportion blame or approval, without thinking that they were responsible in the sense intended here. The parent might be accountable for how the seventeen-year-old child behaves in certain domains. But we would not hold the parent responsible in the sense in which we might hold the child responsible – or might hold the parent responsible, had he or she been the agent. The grounds on which someone can be held accountable are much less demanding than the grounds on which they can be held responsible.

Third, holding someone responsible is distinct from just thinking the person responsible. Holding a person responsible requires thinking that the person is responsible, but it also involves a further component. We think someone responsible when we think that the person satisfies conditions sufficient for being a candidate for blame or approval; we hold them responsible when we go one further step and actually blame or approve. The person who forgives an offender will have to think that the offender was responsible for what was done, and a candidate for blame, else forgiveness would not be in place. But what the forgiveness presumably involves is an eschewal of the blame itself. Again the therapist may believe that a patient was responsible for some deed yet put blame aside. You can imagine the therapist saying: “We both know, of course, that you were fully responsible for that action, but we are not now concerned with blame; our task is to understand why you did it.”

How do we understand the blame that is put aside in such cases or indeed the complementary sort of approval? We need not go into that question, but my assumption is that blaming involves adopting or identifying with the stance of a creditor: someone to whom at least an apology is owed. When other people are involved, adopting such a stance will go with resentment; identifying with the stance will go with indignation.⁴ In one’s own case, adopting or identifying with the stance will go with acknowledging and presumably feeling guilt.⁵

4 The classic source is P. F. Strawson, “Freedom and Resentment,” in *Free Will*, 2nd ed., ed. Gary Watson (Oxford: Oxford University Press, 2003), 72–93.

5 R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge, MA: Harvard University Press, 1994).

Three Conditions for Fitness to Be Held Responsible

Almost everyone is likely to agree that it is often appropriate to hold a group agent causally responsible for certain actions and for the effects of those actions. And again almost everyone will agree that it is often appropriate to hold a group agent accountable for various deeds and effects, for example, those that its employees bring about in the course of doing corporate business. But the question is whether it can also be appropriate to hold a group agent responsible in the richer sense intended here: to think the group is responsible and to adopt or identify with the stance of a creditor – someone to whom a debt is owed.

The focus in the discussion will be on what makes an agent deserve to be held responsible, by our ordinary practices, what constitutes fitness to be held responsible. Since the conditions that make an agent fit to be held responsible will also make the agent fit to be thought responsible, the third of our three distinctions does not matter for the present purposes.

I think that there are three conditions that must be satisfied if someone is fit to be held responsible in a given choice. These conditions correspond to the requirements outlined in some Christian catechisms as conditions necessary and sufficient for a deed to constitute a serious sin. There must have been grave matter, it is said, full knowledge of the guilt, and full consent of the will. The first condition stipulates that the agent faced a morally significant choice; the second that the agent was in a position to see what was at stake; and the third that the choice was truly up to the agent – it was within the domain of the agent's will or control.

Regimenting these ideas a little, we lay down the following three conditions for someone to be fit to be held responsible in a given choice:

Value relevance. – He or she is an autonomous agent and faces a value-relevant choice involving the possibility of doing something good or bad or right or wrong.

Value judgment. – The agent has the understanding and access to evidence required for being able to make judgments about the relative value of such options.

Value sensitivity. – The person has the control necessary for being able to choose between options on the basis of judgments about their value.⁶

These conditions are individually necessary for fitness to be held responsible, since no agent who failed any of them could be reasonably held responsible for an action. Let the agent not face a value-relevant choice, as in the case of trivial decisions. In that case no question of responsibility even arises. Let the agent face a value-relevant choice but not be in a position to make judgments on the values of different options, say because of being denied understanding or evidence on these matters. Again, there will be no basis for holding the person responsible, whether in assigning blame or credit. Finally, let the agent face a value-relevant choice and be able to form a judgment about the value of the options but not have the control required to be able

6 Some of these conditions may be fulfilled without others, leading to familiar scenarios of partial responsibility.

Philip Pettit

to act on that judgment. Here too there will be no basis for regular approval or disapproval; the agent will clearly be beyond the reach of such complaint.⁷

Are the conditions jointly sufficient for fitness to be held responsible, as well as individually necessary? I shall assume that they are since, intuitively, any agent who satisfied them would be a perfectly good candidate for being thought and held responsible.

Holding Responsible and Regulation

Before turning to the issue of how far group agents are fit to be held responsible, it is worth making two further observations. Both bear on the connection, broadly speaking, between holding agents responsible and acting so as to regulate their performance. The one concerns deterrent regulation, as we might call it; the other, developmental regulation.

Deterrent regulation consists in the imposition of sanctions, whether rewards or penalties, with a view to shaping the choices that an agent makes. The most salient form, of course, is the penal regulation of the criminal law, which seeks to shape the behavior of citizens by the threat of legal penalty. Holding someone responsible in the criminal law for having committed a certain offense is distinct from punishing them, say by imposing a fine or prison term or whatever.⁸ My concern here is with how far group agents may be appropriately held responsible, not with how they can be best regulated by penalty or other sanction. I argue that corporate bodies are fit to be held responsible in the same way as individual agents, and this entails that it may therefore be appropriate to make them criminally liable for some things done in their name; they may display a guilty mind, a *mens rea*, as in intentional malice, malice with foresight, negligence, or recklessness. But I say nothing on the practical considerations that are relevant to determining the best sorts of sanctions to impose in criminal law or indeed in the law of tort, the law of contract, or any other branch of jurisprudence. There is a voluminous literature on this subject, and, while the argument certainly makes criminal liability a sensible option to consider with group agents, it does not provide a clear line through the thicket of practical problems that arise in the area.⁹

So much for the distinction between holding agents responsible and regulating them with deterrent sanctions. The other form of regulation that I want to put in the picture is developmental rather than deterrent in character. In order to introduce this idea, think of the way parents often deal with their growing children in domains of behavior where the children may not be fully fit to be held responsible. While

7 Agents with acquired sociopathy may be in the category. See A. L. Roskies, "Are Ethical Judgments Intrinsically Motivational? Lessons from Acquired Sociopathy," *Philosophical Psychology* 16 (2003): 51–66.

8 This is important to recognize because discussions in the theory of criminal justice often run the two together, setting up a false contrast between a backward-looking retributivism and a forward-looking consequentialism. Retributivism should be seen as a doctrine that bears primarily on when someone is fit to be held responsible in criminal law; consequentialism, on how someone who is fit to be held responsible should be sanctioned.

9 See, e.g., Colvin, "Corporate Personality and Criminal Liability"; W. S. Laufer, *Corporate Bodies and Guilty Minds: The Failure of Corporate Criminal Liability* (Chicago: University of Chicago Press, 2006).

recognizing this lack of fitness, parents may yet announce that they will hold children responsible for the good or the bad that they do in the relevant domain, and they may reinforce this attitude with appropriate sanctions. They may allow the teenage son to host a party but insist that they will hold him responsible for any damage done by his friends. Or they may allow the teenage daughter to stay out late but hold her responsible for not missing the last bus. And they may do these things, while being conscious that the children do not yet have the capacities required for reliably achieving the desired results.

Why do this? The most plausible answer is that by treating the children as if they were fit to be held responsible, the parents may help to induce in them the sort of self-awareness and self-regulation that such fitness requires. The practice has a developmental rationale. It makes sense as a way of encouraging in the children those habits that may one day underpin the very fitness to be held responsible that is currently lacking; although the word is not attractive, it has been usefully described as a practice of 'responsibilization'.¹⁰

As holding someone responsible for an action is distinct from punishing or rewarding what is done, thereby seeking a deterrent sort of regulation, so it is distinct from the responsibilizing initiatives in which we might seek to regulate agents developmentally. This is important, as we can imagine adopting initiatives with groups that are designed to achieve a regulatory effect, not to reflect a prior conviction that the groups are truly fit to be held responsible. As children mature and grow in the abilities that are relevant to whether they are fit to be held responsible, so we might think that the same is true of some collectivities. I return to this thought in the conclusion.

II. THE FITNESS OF GROUP AGENTS TO BE HELD RESPONSIBLE: FIRST CONDITION

The Conditions Applied

Applying our three conditions for responsibility to the corporate case, a group agent deserves to be held fully responsible for doing something, we may say, so far as it satisfies these requirements:

Value relevance. – The group is an autonomous agent that faces a significant choice between doing something good or bad or right or wrong.

Value judgment. – The group has the understanding and the access to evidence required for making judgments about the relative value of such options.

Value sensitivity. – The group has the control required for being able to choose between the options on the basis of its judgments about their respective value.

There are two parts to the first condition: the assumption that the group is an autonomous agent, on the one side, and on the other, the assumption that it faces

10 See David Garland, *The Culture of Control: Crime and Social Order in Contemporary Society* (Chicago: University of Chicago Press, 2001); see also Philip Pettit, *A Theory of Freedom: From the Psychology to the Politics of Agency* (Cambridge: Oxford University Press, 2001), chap. 1.

Philip Pettit

a value-relevant choice. The second part does not raise any issues. If a group constitutes an autonomous agent, then it is bound, sooner or later, to face choices involving options that differ in the dimension of value. The question that we have to confront, then, is whether indeed it is possible for groups of the kind envisaged here to constitute autonomous agents.

That question itself breaks down into two. First, is it possible for a group of individual human beings to qualify as an agent, meeting whatever are thought to be the requirements of agency? And if this is possible, then is such a group likely to qualify as an autonomous agent? I defend a positive answer to each question.

Thesis 1: Groups Can Qualify as Agents

A system will constitute an agent, under more or less received analyses, if it forms and reforms action-suited desires for how its environment should be and action-suited beliefs as to how its environment is and if it then acts in such a way that those desires are satisfied according to those beliefs. Action-suited desires will have to be capable of realization by the agent, and they will have to prompt dispositions to realize only scenarios that are consistent with one another. Action-suited beliefs will have to cover the questions that need to be resolved for action, and they will have to represent the environment consistently and in a way that is sensitive to incoming evidence. A system that is to pass as an agent may occasionally fall short of these standards in the formation of its attitudes or may occasionally fail to act as those attitudes require, but it must generally be reliable on those fronts; it must display a robust pattern of attitudinal and behavioral rationality.

This characterization of the requirements of agency allows us to count even simple robotic systems as agents. Imagine the robot that moves on wheels around a tabletop, using robotic arms to change the positions of various little cylinders that are strewn around the table. It scans the table with buglike eyes, and whenever it registers a cylinder on its side, then it moves toward that cylinder and uses its arms to raise it to an upright position. Even a robot of this simple kind can count as an agent. It has a single on-off desire to keep the cylinders upright. It relies on its eyes to register when a cylinder is on its side, where exactly the offending object lies, and how it may be restored to an upright position. And then it acts so as to satisfy its desire according to that set of newly minted beliefs. Or at least it does this in general. When cylinders are near the edge, it may tend to knock them off the table, for example, where the way it moves its arms suggests that that is a failure on its part, not an action prompted by an independent desire.

There are many ways in which the robotic agent might be more complex and might resemble more familiar animal systems. It might form beliefs about objects other than the cylinders and about properties other than those involving their location and orientation. It might form desires for conditions other than having those cylinders assume an upright position. And it might form beliefs and desires relevant to actions at a spatial or temporal remove; it might form intentions to do with other places or times. Such complexities would take it within reach of familiar, nonhuman animals.

Groups can clearly form agents in the sense required under this analysis. They will do so, at least in the normal case, when members act on the shared intention that together they should realize the conditions that ensure agency. They will each intend that together they mimic the performance of a single unified agent. They will each intend to do their bit in the pattern of coordination required for this performance. They will each be motivated to do this by the belief that others intend to do their bit too. And all of that will be above board, as a matter of shared awareness: each will believe that those conditions obtain, believe that each believes this, and so on.¹¹

That the members of a group agent each act on such a shared intention does not mean that they act with equal knowledge or relish or influence, of course. Some may only be aware of the general enterprise and of the particular role required of them; others may also have a sense of the group's detailed strategies. Some may go along with the enterprise for want of a better alternative, others, out of a keen desire for the goals it can accomplish. And some may play small parts, others, very large ones, under the pattern of coordination adopted. Some may even be so alienated that they acquiesce in acting to fulfill intentions that they see others as sharing but do not strictly share themselves.

There are many patterns of coordination – if you like, many constitutions – whereby the members of a group might each be assigned roles in the generation of an action-suited body of desire and belief and in the performance of the actions that it supports. The simplest constitution would have the members assemble and deliberate at regular intervals about the maintenance and development of the group's attitudes and about the actions required; members would vote on the propositions to be accepted as matters of belief or desire and on the actions that they support. Other, more realistic constitutions would depart from that simple model in distributing tasks of attitude formation and action planning among different subgroups or individuals, perhaps with periodic review by the assembly as a whole.

When a group forms a belief or desire or intention on such a pattern, it will do so by appropriately endorsing a corresponding proposition. The group will believe that P, for example, when it or an authorized subgroup or official has considered the proposition and given its assent, according to the accepted constitutional

11 I do not assume that acting on a shared intention is essential for group agency, only that group agency normally involves it, and in any case, I shall concentrate on the cases where it is involved. The line I favor on the analysis of acting on a shared intention is developed further in Philip Pettit and David Schweikard, "Joint Action and Group Agency," *Philosophy of the Social Sciences* 36 (2006): 18–39. Broadly, it follows Michael Bratman, *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999). For other influential approaches see Margaret Gilbert, "Collective Preferences, Obligations, and Rational Choice," *Economics and Philosophy* 17 (2001): 109–20; Raimo Tuomela, *The Importance of Us* (Stanford, CA: Stanford University Press, 1995).

Philip Pettit

formula.¹² The group will form its beliefs, then, by forming on-off judgments over propositions that members put before it. And it will form its desires and specific intentions in a parallel manner by forming on-off preferences over such propositions.¹³

Thesis 2: Groups Can Qualify as Autonomous Agents

There are many ways in which an agent might not be autonomous. An individual agent might fail to be autonomous, for example, through being subject to various compulsions or obsessions or blind spots or *idées fixes* or through being under the unconscious control of another. The salient, if not the only, way in which a group agent might fail to qualify as autonomous is through failing to be an agent that is distinct from the agents who are its members.

It would fail to be autonomous in this manner, arguably, if the attitudes ascribed to it – if you like, the group mind – were just a function of the corresponding attitudes adopted by the members, whether independently or under some scheme of coordination. If the group attitudes were a function of the corresponding individual attitudes, then they would not be novel on any point or proposition; every judgment would reflect corresponding individual judgments, every preference, corresponding individual preferences, and so on. The group system of propositional attitudes would be derivable, proposition by proposition, from individual sets of attitudes so that, given the individual attitudes, the corresponding group attitudes would come for free. Such a group might not be counted as an agent in its own right.

There is a traditional view according to which all group agents are just second-class agents of this kind, not agents with any degree of novelty in the attitudes they embrace. This view is hard to defend in relation to complex, organized entities, and, perhaps for that reason, the defense is usually undertaken with reference to the more plausible case provided by participatory agents in which each member plays more or less the same sort of role. The suggestion is that seeing group agency present in the way such members act is the product of an illusion. It is like looking at rush hour traffic from on high and imagining that the lines of cars are entities that each act out a mind of their own.

Anthony Quinton defends a version of the traditional view. “To ascribe mental predicates to a group is always an indirect way of ascribing such predicates to its members. With such mental states as beliefs and attitudes, the ascriptions are of

12 This applies to the explicit processes of belief formation that must figure in any mental life. Some propositions will have to receive the assent of the group, however, by way of being presupposed in other more explicit patterns. That, in effect, is the lesson of Lewis Carroll, “What the Tortoise Said to Achilles,” *Mind* 4 (1895): 278–80.

13 While the propositions endorsed may be probabilistic, the possibility ruled out is that members might put degrees of belief and desire together, mainlining one another’s minds and generating a system of probability and utility for the group as a whole. Assuming that people do each have fine-grained degrees of belief and desire, it is not clear how they could even know what they are, let alone communicate them to one another in the mainlining fashion; such degrees will show up in behavior, particularly in dispositions to accept various gambles, but need not be available to introspection. See Gilbert Harman, *Change in View* (Cambridge, MA: MIT Press, 1986).

Table 1: *Matrix 1*

	P?	Q?	R?	P&Q&R?
A	No	Yes	Yes	No
B	Yes	No	Yes	No
C	Yes	Yes	No	No
A, B, and C	Yes	Yes	Yes	No

what I have called a summative kind. To say that the industrial working class is determined to resist anti-trade union laws is to say that all or most industrial workers are so minded.”¹⁴ The idea is that the attitudes ascribed to a group agent are just a logical function – here a majoritarian function – of the corresponding attitudes held by the members and that it is an illusion, therefore, to imagine that the group is an agent that exists and operates side by side with its members, enacting a novel and distinct system of attitudes. Were we counting the agents present, we would have to count the individual members as distinct agents but not the group agent that they constitute.

This debunking view of group agency is demonstrably mistaken, even with participatory groups. “Discursive dilemmas” show that it is wrong to think that as a general rule group attitudes can be a majoritarian function of member attitudes.¹⁵ And associated impossibility theorems explain why it is wrong to think that they can be any sort of function, majoritarian or nonmajoritarian, of such attitudes. The lesson is that if a set of individuals are to constitute a group agent, even in the participatory manner, then they cannot rely on the group attitudes being formed on the basis of the corresponding attitudes among members.

Consider a group of three agents, A, B, and C, that tries to conform to majority voting in the formation of its judgments; a similar lesson might be derived for preferences. Imagine that under the pressure of decision and action, they have to form judgments, now on whether P, now on whether Q, now on whether R, and yet again on whether P&Q&R. All but A may vote for P; all but B, for Q; all but C, for R; and, consequently, none for P&Q&R: each will reject it because of rejecting one conjunct. Under a majoritarian arrangement, then, these votes would have the group holding that P, that Q, that R, and, inconsistently, that not P&Q&R. The situation is summarized in matrix 1.

There is a permanent threat of such inconsistency with any corporate entity that tries to do its business by majority voting. Such an entity will exist over time, con-

14 Anthony Quinton, “Social Objects,” *Proceedings of the Aristotelian Society* 75 (1975).
15 See Pettit, *Theory of Freedom*, chap. 5, and “Groups with Minds of Their Own,” in *Socializing Metaphysics*, ed. Frederick F. Schmitt (New York: Rowan & Littlefield, 2003), 129–66. The idea of the discursive dilemma is a generalization of the legal idea of a doctrinal paradox. See L. A. Kornhauser and L. G. Sager, “The One and the Many: Adjudication in Collegial Courts,” *California Law Review* 81 (1993): 1–59. For an overview of the topic and of other issues, see Christian List, “The Discursive Dilemma and Public Reason,” *Ethics* 116 (2006): 362–402.

Philip Pettit

fronting a range of issues: now whether P, now whether Q, now whether R, and so on. Sooner or later, under plausible assumptions, it is very likely to face an issue where majority judgment is inconsistent with the judgments already on record.¹⁶ It might confront the issue of whether P&Q&R at a later time, for example, although not necessarily under such a perspicuous representation.

What is the A, B, and C group to do in such a case, assuming that members each intend that they act together as a single group agent? By hypothesis, they need to have a complete set of views on the issues involved; they cannot suspend judgment in any case, given that the issues are closely connected with action. Yet they cannot live with the inconsistent set of views that majority voting would give them, for such irrational views are liable to point to inconsistent sets of actions and to present the group as an agent with which it makes no sense to form contracts or enter other relationships. The only solution for the members of the group will be to embrace a practice or constitution that allows them to ensure that the body of attitudes they accept and enact in the group's name is internally consistent.

But this solution will ensure that the group attitudes cease to be a majoritarian function of the member attitudes and will cast it to that extent as an autonomous subject. The constitution will have to allow the members as a group to hold that not P, not Q, or not R, contrary to one or another majority, or that P&Q&R, contrary to every one of the members. There will be at least one issue where the group's judgment fails to coincide with a corresponding judgment on the part of a majority of members. The group judgment may even be a judgment that every member rejects, as in the case where the members decide, or the constitution they adopt automatically ensures, that as a group they endorse the judgment that P&Q&R.

The permanent possibility of such a problem indicates that if a group is to form judgments on a connected set of issues, and to do so consistently, then there is no guarantee that every judgment formed can be a majoritarian function of the corresponding judgments on the part of individuals. On the contrary, the need to guard against the sort of difficulty illustrated entails that the group will have to follow a nonmajoritarian procedure. The members will have to create a group agent that comes apart in that manner from the way that they are individually disposed.

The position of such a group agent in relation to the members will be a little like that which might obtain between an individual person and those experts he or she consults as advisers, generally accepting any view that a majority supports. The advisee will not be able to go along slavishly with majority opinions since, as the discursive dilemma shows, these can be inconsistent with one another. Presumably the only satisfactory line will be to go along with majority opinion subject to the constraint of not allowing internal inconsistency. That policy will not dictate a uniquely best set of judgments to adopt, but it will at least ensure that the advisee remains an independent center of judgment, not just a voice box of those

16 On this likelihood see List, "Discursive Dilemma and Public Reason."

consulted.¹⁷ In precisely the same way the only satisfactory line for a participatory group may be to go along with the majority view of its members on any issue but to do so only under the constraint of ensuring that this does not lead to inconsistency. Thus, like the advisee in relation to the expert advisers, it will be heavily influenced by the attitudes of its members but will still count as a distinct attitudinal center and a distinct agent.

Although a group agent's attitudes cannot in general be a majoritarian function of member attitudes, it may be that they are still functionally related to corresponding individual attitudes and that a nonmajoritarian counterpart of Quinton's debunking view is sound. But such a counterpart view is ruled out by impossibility theorems associated with the discursive dilemma.¹⁸ Broadly, these results show that any group that is robustly disposed to form consistent and complete judgments over certain connected propositions, as any rational group agent must be disposed to do, will have to be disposed to form judgments that are not a function, majoritarian or otherwise, of the corresponding judgments of its members: this, at any rate, so long as its judgments depend on the judgments of more than one individual.

Assume, for example, that a group of individuals has to find a procedure for deriving a set of judgments on certain logically connected propositions. Take any procedure that makes the judgments of the corporate body depend on the judgments of more than one individual, thereby ruling out dictatorship of any kind. Let this procedure work for any consistent sets of input judgments, enabling the group to produce complete and consistent judgments; in other words, let it robustly guarantee the rationality of the group. No procedure that satisfies conditions of these kinds will identify a rule or function – majoritarian or nonmajoritarian – whereby the corporate judgment on every issue automatically derives from the votes and judgments of members on that issue.¹⁹

The lesson of this and related results is that the attitudes of reliable group agents cannot be a majoritarian or nonmajoritarian function of the corresponding attitudes among individuals, and, as follows under some plausible assumptions, that they cannot even be fixed by a mix of such functions. Does this mean that group

17 Philip Pettit, "When to Defer to a Majority - and When Not," *Analysis* 66 (2006): 179–87.

18 The initial theorem, explicitly focused on generalizing the discursive dilemma, is proved in Christian List and Philip Pettit, "Aggregating Sets of Judgments: An Impossibility Result," *Economics and Philosophy* 18 (2002): 89–110. For a more recent theorem see Franz Dietrich and Christian List, "Arrow's Theorem in Judgment Aggregation," *Social Choice and Welfare* (forthcoming); this relates closely to results in Marc Pauly and Martin Van Hees, "Logical Constraints on Judgment Aggregation," *Journal of Philosophical Logic* (forthcoming).

19 This is the Dietrich-List result. The List-Pettit result varies a little. Assume that a group of individuals has to find a procedure for deriving a set of judgments on propositions that are logically connected, although in a slightly different sense from that in the other case. Take any procedure that treats every individual voter equally (anonymity), giving no one the status of a dictator; that is designed to work for any consistent sets of input judgments (universal domain); and that enables a group to produce complete and consistent judgments over ranges of connected issues. No procedure that satisfies conditions of roughly these kinds will identify a rule or function – majoritarian or nonmajoritarian – whereby the corporate judgment on every issue can be derived from the votes and judgments of members on that issue.

Philip Pettit

agents are autonomous in relation to their members? I shall assume that it does. Autonomy is intuitively guaranteed by the fact that on one or more issues the judgment of the group will have to be functionally independent of the corresponding member judgments, so that its intentional attitudes as a whole are most saliently unified by being, precisely, the attitudes of the group. This autonomy may be surprising, but it is not mysterious. While group attitudes are not functions of the corresponding attitudes of individual members, they are produced by those individuals, and they derive all their matter and energy from what individuals supply.²⁰

III. THE FITNESS OF GROUP AGENTS TO BE HELD RESPONSIBLE: SECOND CONDITION

Human and Simpler Agents

We have seen that corporate entities can be bona fide agents and agents that are relatively autonomous in relation to their members. They can qualify as autonomous agents, in the sense of the first condition for fitness to be held responsible, and agents who may confront evaluatively significant choices, in which matters of good and bad and right and wrong are at stake. What, then, of the second condition? How likely is it that group agents will be able to make judgments as to the relative value of the options that they face in such a choice? How likely is it that such an agent will be able to understand value judgments about options and to access evidence on the relative value of the options it faces in a certain choice?

The crucial issue is whether a group agent will be able to understand value judgments about options, for if it can do this, then there are bound to be some circumstances in which it will have access to evidence about the relative value of the options it faces.²¹ A group agent will be able to understand value judgments, just in case the propositions over which it is able to form attitudes, in particular judgments, include propositions bearing on the relative value of the options it faces. So the question is whether it is plausible to claim that the propositions in the domain of a group's attitude formation might not include propositions of that type.

There is nothing incoherent in the idea that group agents might not be able to form judgments over such propositions, for certain systems may count as intentional agents in the absence of this ability. As we saw earlier, intentional agents will have to be reliably disposed to act in a way that satisfies their desires according to their

20 Christian List and Philip Pettit, "Group Agency and Supervenience," *Southern Journal of Philosophy*, suppl., ed. Murray Spindel 44 (2006): 85–104. For an argument that the attitudes of reliable group agents cannot be generated without group-level reasoning, so that autonomy is supported in yet a stronger fashion, see Philip Pettit, "Personal and Sub-personal Reason: The Case of Groups," *Dialectica* 61 (2007), forthcoming.

21 We sometimes judge that ignorance of evidence was invincible and blameless; sometimes, that it was not: we rule that, however ignorant, the agent still had access to the evidence. While it is hard to say what principles should govern such judgments of evidential access, we need not dwell on the question here, since it arises in the individual case as well as in the case of group agents. See G. Rosen, "Scepticism about Moral Responsibility," *Philosophical Perspectives* 18 (2004): 295–313.

beliefs. But that disposition is consistent with their not being able to hold any beliefs about the desirability – the desire-worthiness – of options. For all that the notion of intentionality requires, the intentional agent need only have beliefs that serve to channel its actions toward the satisfaction of its desires. It need not have any critical or evaluative beliefs: that is, any beliefs according to which a certain goal or a certain action ought to be desired, whether from a rational, prudential, or moral point of view.

It is clear that a system as simple as our little robot will certainly not be able to form evaluative beliefs, given how restricted is its domain of attention. But equally, there is little or no evidence that more complex, nonhuman animals can form such beliefs either. That is surely the reason why we have little or no inclination to hold such creatures responsible for what they do. They are not fit to be held responsible because they do not make the value judgments that supply the benchmarks of responsibility.

The options considered in a choice involve possibilities of action and outcome that abstract away from the particular manner in which they might materialize. They are propositions corresponding to the sentences that describe the options: the propositions expressed by “I do X,” “I do Y,” “I do Z,” and the like. When we human beings reason in practical mode about what to do, then we form beliefs about those propositions, say to the effect that by realizing one such possibility we will realize another and that the possibilities involved have this or that degree of appeal, rational, prudential, or moral. We form metapositional beliefs akin to those formed when we reason theoretically, asking ourselves about relations of consistency among propositions or about their inductive support.

Simpler intentional systems may form beliefs in propositions about environmental objects and their properties – about cylinders, for example, and their orientation. But it seems extravagant to think that they can also form beliefs in propositions about propositions. We can do this, at least in part, because we have access to our distinctive kind of language. We can use a sentence such as “I do X” or “A life is saved” or “Joan is happy,” not just to report that result but to exemplify the possibility it expresses and then to hold up that possibility as an object of attention and to ask about its properties.²² This is something that lies beyond the reach of nonhuman animals and artifacts. The dog may hear the noise at the gate, prick up its ears in attention, and ask the question, in effect, of whether dinner is being served. When we deliberate about what to do, we go through a similar routine at a more abstract level. We attend to abstract propositions or possibilities, not just concrete objects, and we ask questions as to whether a given way that things may be – a way we may make them to be – is good or right or better than other alternatives.

22 Philip Pettit, *The Common Mind: An Essay on Psychology, Society, and Politics* (1993; repr., New York: Oxford University Press, 1996), chap. 2.

Philip Pettit

Group Agents

Might group agents, like nonhuman animals, be incapable of making judgments about the relative value of the options they face? Or are they like individual human beings, who certainly can make such judgments?

A group will form a judgment or other attitude over a certain proposition when the proposition is presented for consideration and the group takes whatever steps are prescribed in the constitution for endorsing it; these steps may involve a vote in the committee of the whole, a vote in an authorized subgroup, or the determination of an appointed official. Thus, a group will be able to form a judgment over any proposition that members are capable of presenting for consideration and of adjudicating by means of a vote or something of the kind. But, since the members of any group will be able to do this in their individual lives, they will certainly be able to do it when acting within the group. They will be able to present evaluative, option-related propositions for group consideration and will be able to take a part in deciding them. It follows, then, that group agents must be able to form value judgments about the options they face in any choice and that the second condition for group responsibility is bound to be fulfilled.

The ability to do something, of course, may be a more or less remote or proximate ability. It may be the ability of an agent to perform an action, just like that, or it may be the ability of the agent to take steps that will make it possible eventually to perform the action. The argument just given only establishes that group agents must have the remote ability to form judgments about the relative value of options they face; it does not mean that they have the proximate ability to do so. Thus, the constitution of a group may restrict its agenda to propositions of a nonevaluative kind, in which case the members will be able to present evaluative propositions for consideration and require that they be adjudicated, only so far as they are able to get that constitution changed. To be sure, members will always be able to change the constitution and to usher in evaluative judgment, but the constitutional restriction may make that difficult and ensure that the exercise of the evaluative ability is only a remote prospect.

This consideration need not be disturbing from our point of view, however, for two reasons. The first is that few group agents are likely to impose constitutional restrictions against forming evaluations of the options they face, even if they are not in the habit of making such appraisals. And the second is that it would seem to be a serious design fault, at least from the viewpoint of society as a whole, to allow any group agents to avoid making evaluations of this kind. Why should any group of individuals be allowed to incorporate under a constitution that deprives it of the ability to evaluate its options and that ensures, therefore, that it will not be fit to be held responsible for what it chooses?

IV. THE FITNESS OF GROUP AGENTS TO BE HELD RESPONSIBLE: THIRD CONDITION

A Problem

The question raised by the third condition is whether a group agent, deliberating as a group, is in value-sensitive control of which option to take, being able to go where its evaluative judgments lead. Given that the group agent is able to form judgments on the relative value of different options, the question is whether it can make choices in response to those judgments: whether it can be duly sensitive to the reasons that the judgments provide. This is an open question, since even individual agents may be subject to pathologies and weaknesses that disable them from reliably acting on the value judgments that they make.

The notion of value-sensitive or, more generally, reason-sensitive control needs analysis in any full theory of agency. Since it is required in the theory of individual agency as much as in the theory of group agency, however, we need not try to provide an analysis here.²³ The challenge before us is not to explain what reason-sensitive control is but rather to show that there is no particular reason why reason-sensitive control, whatever it involves, shouldn't be instantiated in a group agent. I shall proceed on the assumption that it is indeed instantiated in individual agents, and so the focus will be on whether there is any difference between individual and group agents that explains why group agents should lack it.

There is one difference in particular that may seem to stand in the way of ascribing reason-sensitive control to group agents. Whatever a group does is done by individual members on behalf of the group and is done intentionally by those individuals. But in that case, it would appear that the group cannot be in intentional, reason-sensitive control of what it does. Such control may seem to lie entirely with those who act in its name. The group may be able to make value judgments about the options it faces, but when it comes to acting, it may find itself unable to give those judgments any effective role. It may be that the members who act in its name have exclusive control over what is done and exclusive responsibility for doing it.

This difficulty has long been registered in the philosophical and theological tradition. In 1246, Pope Innocent IV argued that a corporate body, or *universitas*, cannot be excommunicated, being only a fictional person, not a real one.²⁴ That argument was endorsed by Aquinas on the basis of what one commentator describes as "the old and solid truth that only individuals can act and, more especially, that only individuals can commit a delict and become guilty."²⁵ Building on this axiom, St. Thomas argued in support of the Pope's line. Anything done by a group agent, so

23 Philip Pettit and Michael Smith, "Freedom in Belief and Desire," *Journal of Philosophy* 93 (1996): 429–49, reprinted in Frank Jackson, Philip Pettit, and Michael Smith, *Mind, Morality, and Explanation* (Oxford: Oxford University Press, 2004); Pettit, *Theory of Freedom*.

24 E. H. Kantorowicz, *The King's Two Bodies: A Study in Mediaeval Political Theology* (Princeton, NJ: Princeton University Press, 1997), 305–6.

25 Th. Eschmann, "Studies on the Notion of Society in St. Thomas Aquinas: St. Thomas and the Decretal of Innocent IV, *Romana Ecclesia; Ceterum*," *Medieval Studies* 8 (1946): 35. On Aquinas's argument, see 11.

Philip Pettit

the argument went, is done either by some members or by all. If it is done by some members, the fault lies with them. If it is done by all, then the fault is divided up among members as individual agents; it does not belong in any independent sense to the body corporate.

Put abstractly, the problem we face is generated by the following argument:

Whatever a group agent does is done by individual agents. Individuals are in reason-sensitive control of anything that they do and so in control of anything they do in acting for a group. One and the same action cannot be subject both to the reason-sensitive control of the group agent and to the reason-sensitive control of one or more individuals.

Hence, the group agent cannot be in reason-sensitive control of what it does; such control will always rest with the individuals who act for the group.

This argument is clearly valid so that if we are to reject its conclusion, we must find fault with one or more of the three premises. But we have been proceeding on the assumption that the first two premises are sound, as indeed they intuitively are. Thus, the question is whether we have to accept the third, crucial premise, which denies the possibility that reason-sensitive control of an action might be exercised at once by the group and by the member or members who act for the group.

At first it may seem that that premise must be false. Those who act for group agents are typically described as acting on the instructions of the group or by the commission of the group. But in that case it seems that both the group and the members can each have control over the action: the group as the agent that gives instructions on what is to be done and the members as the agents who carry out those instructions. This line won't work, however. Just as anything the group does is done by members, so any instructions that the group gives are given by one or more members. And so the problem recurs one stage earlier. How can the group be in reason-sensitive control of the instructions given, if those instructions are already under the reason-sensitive control of the members who issue them?

One way to sharpen this problem is to notice that things would be very different if group agents resembled swarms of bees, say, swarms in pursuit of a new nesting location.²⁶ We can imagine that when bees behave as groups, pursuing swarm-level goals according to swarm-level representations, the individual bees may just be responding in a quite mechanical, attitudinally unmediated way to chemical and other signals. The swarm may relate to the bees in the way an intelligent organism relates to its cells: it may be intentionally minded, as an organism may be intentionally minded, while the bees, like the cells, are more or less mindless automatons. Under such a picture, there will be no problem with the idea that what the swarm

26 T. D. Seeley, "Decision Making in Superorganisms: How Collective Wisdom Arises from the Poorly Informed Masses," in *Bounded Rationality: The Adaptive Toolbox*, ed. Gerd Gigerenzer and Reinhard Selten (Cambridge, MA: MIT Press, 2001). For background see R.A. Wilson, "Group-Level Cognition," *Philosophy of Science* 68 (2001): 262–73.

intentionally controls – the control, of course, won't be reason-sensitive control – is controlled only by that entity; the individual bees won't act intentionally and so cannot exercise intentional control.

The problem in the human case is that the same sort of control we ascribe to the group agent is exercised also by individual members. Given that corporate agents only act intentionally via the intentional, reason-sensitive actions of their members, can they really have the sort of control that would be required for responsibility? Do they have the capacity to give effect in their choices to the value judgments they can make over the options they face?

A General Perspective

The general issue in the background of the problem raised is one that arises in nonagential contexts as well as agential. It is the problem of how there can be higher-level and lower-level factors that are causally relevant to one and the same event; and this, despite the fact that neither factor causes the other, and neither combines with the other as part of a larger cause. The problem in particular is to explain how lower-level causes can fail to exclude the possibility of higher-level causation.²⁷ It may be useful to look at that more general problem and at a plausible way of dealing with it, before returning to the corporate case.

Consider a natural process in which water in a closed flask is brought to the boil and, as a consequence, the flask breaks. Let us assume that what happens in the process is that as the water boils – as the mean motion of the constituent molecules reaches a suitable level – it becomes likely to the point of near inevitability that some molecule will have a position and momentum sufficient to break a molecular bond in the surface of the flask and that this in fact happens, leading to the collapse of the flask. What causes the flask to break in such a case?

At one level the molecule that actually triggers the break in the surface is the cause of the collapse. But the fact that the water is boiling is also causally relevant to the event. The boiling temperature of the water consists in the mean molecular motion being at such and such a level and so is constituted by the motion of the triggering molecule together with the motion of the other molecules. Thus, it cannot be a cause of the motion of the triggering molecule and cannot cooperate with that motion as part of a larger cause. But it is still causally relevant.

Its relevance consists in the fact that given its presence – given that the water is boiling – it is more or less inevitable that there will be some constituent molecule, maybe this or maybe that, that has a position and a momentum sufficient to induce a crack in the surface of the flask.

The relationship between the causally relevant temperature and the causally relevant molecule might be described in terms of a metaphor from computing.²⁸ The

27 Jaegwon Kim, *Mind in a Physical World* (Cambridge, MA: MIT Press, 2000).

28 Frank Jackson and Philip Pettit, "Functionalism and Broad Content," *Mind* 97 (1988): 381–400, and "Program Explanation: A General Perspective," *Analysis* 50 (1990): 107–17. Both are reprinted in Jackson, Pettit, and Smith, *Mind, Morality, and Explanation*.

Philip Pettit

higher-level event – the water being at the boiling point – programs for the collapse of the flask, and the lower-level event implements that program by actually producing the break. The facts involved, described more prosaically, are these. First, the higher-level event may be realized in many different ways, with the number, positions, and momenta of the constituent molecules varying within the constraint of maintaining such and such a mean level of motion. Second, no matter how the higher-level event is realized – no matter how the relevant molecules and motion are distributed – it is almost certain to involve a molecule that has a position and momentum sufficient to break the flask. And, third, the way it is actually realized does have a molecule active in that role.

Given the fulfillment of such conditions, we can say that the water being at boiling temperature programs for the breaking of the flask, whereas the molecule behaving as it does in such and such a vicinity implements that program; it plays the immediate productive role. Both programming and implementing are ways, intuitively, of being causally relevant, and so it makes sense, depending on context, to invoke one or the other in causal explanation of the effect; information about either antecedent, higher level or lower level, will amount to important information about the causal history of the event.

Resolving the Problem

The analogy with the water gives us a helpful angle on our problem with the reason-sensitive control of group agents over what course of action they take. Suppose that the member of a group does something in the corporate name, exercising intentional and reason-sensitive control in the usual manner. Is there any sense in which the corporate organization can also exercise such control over what is done? The answer is that it can share in that control, so far as it relates as a programming factor to the implementing factor represented by the active individual.

The temperature of the water controls for the breaking of the flask, so far as it ensures, more or less, that there will be some molecule, maybe this or maybe that, which has a momentum and position sufficient to trigger a breaking; the molecule itself controls for the breaking, so far as it ensures that this particular crack materializes in the surface of the flask. Things may be perfectly analogous in the case of the group agent, except that the mode of control is different. The group may control in a reason-sensitive way for the performance of a certain action by some members, maybe these or maybe those. It will do this, by maintaining a constitution for the formation and enactment of its attitudes, arranging things so that some individual or individuals are identified as the agents to perform a required task, and other individuals are identified as agents to ensure that should the performers fail, there will be others to take their place as backups. Consistently with this group-level control, however, those who enact the required performance will also control in a reason-sensitive way for what is done; they will control for the fact that it is they and not others who actually carry it out.

Under this story, the group agent will be fit to be held responsible for ensuring, or more or less ensuring, that one or more of its members will perform in the relevant manner. The contributing member or members of the group will not be absolved

thereby, however, of their own responsibility. For, other things being equal, they will each be fit to be held responsible for the fact that it is they who actually help to get the action performed. The members will have responsibility as enactors of the corporate deed, so far as they could have refused to play that part and didn't. The group agent as a whole will have responsibility as the source of that deed: the ultimate, reason-sensitive planner at its origin.

This line of thought resolves the problem that seemed to block the possibility that a group agent might satisfy the third condition necessary for fitness to be held responsible. A group agent, so it now transpires, is as fit as any individual human being to be held responsible for what it does. The members of such an agent combine to form a single agent, capable of making a judgment on what is good and bad and right and wrong and capable of ensuring, in awareness and sensitivity to the demands of value, that one or another line is taken in the choices that it faces. The individuals who act for the group on any occasion will have to answer, of course, for what they do. But the corporate entity will also have to answer for what it ensures will be done, drawing on the resources provided by such members. It has all the capacities of reason-sensitive agency necessary to make this demand entirely plausible.

V. INDIVIDUAL AND CORPORATE RESPONSIBILITY

Is There a Point in Holding Group Agents Responsible?

The argument so far shows that group agents can meet the three conditions for being thought and held responsible. But it is one thing to say that there is no bar to holding group agents responsible, given that they can satisfy the conditions reviewed. It is quite another thing to argue that there is a point to this exercise. Someone might maintain, after all, that so long as we hold members responsible for their individual contributions to the doings of a group agent, there will be no practical gain, and there may even be a disadvantage, in going on to hold the group as a whole responsible as well.

The challenge gives us the question with which we shall be concerned in this section. How far does the fact of holding individuals responsible for what is done in a group's name make it unnecessary or redundant, even perhaps counterproductive, to hold the group corporately responsible as well?

There are three sorts of responsibility that individuals may bear in relation to the behavior of a group agent, and it will be useful, as a preliminary, to identify the sort of individual responsibility that gives rise to the question. Individuals may be held responsible for what a group does as designers of a group agent, as members of the group, or as those who enact its wishes. Or of course they may bear responsibility under two or even three of these headings. I argue that the question we are concerned with involves the responsibility of individuals as enactors, not in any other guise.

Philip Pettit

Individuals will be responsible as designers of a group, so far as they determine the technology that the group uses for forming its preferences and judgments and making its decisions. The founders of any corporate entity, be it a church body, a political party, or a commercial organization, necessarily bear some responsibility for how the group functions in the future as a result of the decisions that they make. This sort of individual-level responsibility is not of much relevance to our question, however, since it leaves in place the responsibility of the group for doing what the designers made it possible to do. The responsibility of the designers in relation to the later performance of the corporate entity is like the responsibility of parents in relation to the later behavior of their children. As parents may have laid down formative habits in their children, whether for good or bad, so the designers will have laid down formative routines in a group they shape. But as the impact of parents in this regard normally does little to reduce the fitness of the children to be held responsible for their actions, so the same is true of the impact of designers on the behavior of a group.

Member responsibility, as I understand it here, is the responsibility that individuals have as the members of a group agent that does good or bad. This sort of responsibility is also irrelevant from the point of view of the question raised since it is derivative from the responsibility of the group agent, not a form of responsibility that might compete with it. Individuals have member responsibility for what a group does, so far as it is their group agent, their church or association or company, which produces that result. As continuing, nonprotesting members, they will be complicit in the group performance.²⁹

The third way in which individuals are fit to be held responsible for anything that a group does is as enactors of the group's plans. Other things being equal, as we saw, enactors will be responsible for anything they do in the group's name, to the extent that they could have refused to play that part. This responsibility is consistent with the group being responsible as a whole for ensuring that someone plays that part. But the question is whether there is any point in alleging the group's fitness to be held responsible for an action, given that this enactor responsibility will accrue to individual members. When we have identified the individuals who bear this responsibility, won't it be at least redundant to go on and argue that the group as a whole is fit to be held responsible as well? And mightn't it even be counterproductive, relieving the individual enactors of some of the responsibility that, intuitively, they should be forced to carry? It is in the context of such enactor responsibility that the question has bite.

The question is particularly challenging once we realize that the category of enactor ought to include negative as well as positive enactors. The negative enactors with any group action will be those who could have blocked or inhibited or at least challenged the initiatives of the positive enactors and omitted to do so.³⁰ Once we have ascribed full individual responsibility to the positive and negative enactors of a

29 Christopher Kutz, *Complicity: Ethics and Law for a Collective Age* (Cambridge: Cambridge University Press, 2001).

30 See Keith Graham, "Morality, Individuals, and Collectives," in *Moral Philosophy and Contemporary Problems*, ed. J. D. G. Evans (Cambridge: Cambridge University Press, 1987), 12

group action, will there be any point in going on to assign corporate responsibility? That is the question we face.

Why Group Agents as Well as Members Should Be Held Responsible

I hold that even when all the relevant enactors in a group action have been identified and held responsible, still it may be important to hold the group agent responsible as well. The reason for this, very simply, is that it is possible to have a situation in which there is ground for holding the group agent responsible, given that it satisfies the three conditions listed, but not the same ground for holding individual enactors responsible. The responsibility of enactors may leave a deficit in the accounting books, and the only possible way to guard against this may be to allow for the corporate responsibility of the group in the name of which they act.

In order to introduce the argument for this view, it may be useful to consider first the more familiar case of individuals who do not incorporate as a group agent but combine to have a common effect or, on a one-off basis, form a joint intention and perform a joint action. Many people have argued that such unincorporated collections of agents may act in a way that predictably brings about bad results, without the members of the collection being individually or distributively culpable, or at least not in a serious measure.³¹ It may be that the individuals are ignorant, and blamelessly ignorant, of the harm that they together bring about. It may be that they each take themselves not to make a pivotal difference to a harm done, as with the firing squad in which members each treat the behavior of the others as fixed. It may be that they take themselves to make a difference, but not the right sort of difference, in particular not the sort that increases the harm; for example, each driver in a group of dangerously speeding cars may see that he or she dare not slow down, for fear of making a bad outcome worse.³² Or it may be that while each is aware of the harm done, and aware of making a difference, even the right sort of difference, still they each act under such felt pressure – perhaps pressure from one another – that they cannot be held fully responsible for the contribution they make to a bad outcome; they can each reasonably argue that the circumstances mitigate the degree of their personal responsibility.³³

As individuals, the members of an unincorporated collection are agents, but in cases of these kinds they are not fully culpable agents, since each has at least a partial excuse for his or her behavior. But neither can the members be held collectively, as distinct from individually, responsible, that is, responsible as a collection or group.

31 But see Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), pt. 1.

32 Frank Jackson, "Group Morality," in *Metaphysics and Morality: Essays in Honour of J. J. C. Smart*, ed. Philip Pettit, Richard Sylvan, and Jean Norman (Oxford: Blackwell, 1987), 91–110. In the firing-squad case, there is a weak suboptimal equilibrium from which no one can unilaterally depart with moral benefit. In the speeding-car case, there is a strong suboptimal moral equilibrium from which no one can unilaterally depart without doing moral harm.

33 P. A. Werhane and R. E. Freeman, "Corporate Responsibility," in *The Oxford Handbook of Practical Ethics*, ed. H. LaFolette (Oxford: Oxford University Press, 2003), 523–24.

Philip Pettit

For the collection that they constitute, being unincorporated, is not an agent. It fails the first condition for fitness to be held responsible.³⁴

The considerations that would make the members of a collection less than fully responsible as individuals for a collective effect can have the same impact on the responsibility that attaches to the members of a corporate agent. Certainly the members of an unincorporated collection may be less than fully responsible as individuals, so far as they are blamelessly ignorant of any harm that is collectively done; reasonably believe that they won't make a difference to the harm done, or at least not the right sort of difference; and act under a sense of duress or pressure from others. But the very same can be true of the members of a group agent. Thus, a group agent can act to bring about some harm, or indeed some good, without any of its members being fully fit to be held responsible for his or her contribution to that result.

Shortfalls of individual responsibility have a distressing aspect in the case of the unincorporated collection, since they mean that although the individuals do something bad together, there is no one to hold responsible. But the failures of individual responsibility in the case of the incorporated group may leave us with someone to hold responsible: the group agent itself. That group, by the argument of the last few sections, is an agent that may be fit to be held responsible. And the fact that enactor responsibility can fall short in the ways illustrated means that there may be very good reason to hold the group responsible in addition to holding the enactors responsible. We can hold the enactors responsible, so far as their circumstances allow this, for the harm that their voluntary acts and omissions occasioned. We can hold the corporate entity responsible for the harm that it arranges to have done, given the decisions it licenses and the constitution by which it channels those decisions.

There will be reason to hold group agents responsible in general, not just in cases where there is evidence of a shortfall in individual fitness to be held responsible. This is for two reasons. First, there is always likely to be a shortfall in enactor responsibility, and it is important to guard against that possibility. The possibility is particularly salient, given our earlier story about the relative autonomy with which corporate attitudes may form; this is likely in many cases to leave individuals blamelessly ignorant of how their contributions will affect corporate performance. And second, the failure to impose a regime of corporate responsibility can expose individuals to a perverse incentive. Let human beings operate outside such a regime, and they will be able to incorporate, so as to achieve a certain bad and self-serving effect, while arranging things so that none of them can be held fully responsible for

34 See Virginia Held, "Can a Random Collection of Individuals Be Morally Responsible?" *Journal of Philosophy* 67 (1970): 471–81. Matthew Braham and M. J. Holler ("Distributing Causal Responsibility in Collectivities," in their *Economics, Rational Choice, and Normative Philosophy* [London: Routledge, 2007]) argue that it is always possible to distribute causal responsibility among the members of a collectivity who together bring about some result. That may be so, but my concern here is with moral responsibility, not causal.

Table 2: Matrix 2

	Serious Danger?	Effective Measure?	Bearable Loss?	Pay Sacrifice?
A	No	Yes	Yes	No
B	Yes	No	Yes	No
C	Yes	Yes	No	No

what is done. Things could be fixed so that the individuals are protected by excusing or exonerating considerations of the kind that we rehearsed earlier.

I conclude that not only is it going to be possible to hold group agents responsible, as their satisfaction of our three conditions ensures, it is also likely to be desirable. Let group agents be freed from the burden of being held responsible, and the door will open to abuses: there will be cases where no one is held responsible for actions that are manifestly matters of agential responsibility. The regime I envisage would hold individual enactors responsible for any harm that they might have refused to do and didn't. And it would hold the corporate agent responsible for having organized things so that such harm was likely or inevitable.³⁵

An Example

Imagine a commercial company that is owned and effectively run by its employees in the manner of a participatory organization. And now imagine that it faces an issue about whether to forgo a pay rise in order to spend the money thereby saved on introducing a workplace safety measure, say, a guard against electrocution.³⁶ Let us suppose that the employees have agreed to make the decision on the basis of the majority view on three separable questions: first, whether there is a serious danger of electrocution, by some agreed benchmark; second, whether the safety measure that a pay sacrifice would buy is likely to be effective against the purported danger, again by an agreed benchmark; and third, whether the pay sacrifice involves an intuitively bearable loss for individual members. If a majority thinks that the danger is sufficiently serious, the safety measure sufficiently effective, and the pay sacrifice sufficiently bearable, the pay sacrifice will go through; otherwise, it will not.

Imagine now that after appropriate dialogue and deliberation the employees make individual judgments on the four propositions in the pattern illustrated by the following matrix for a sample of three workers; the structure will be familiar already from the case with P, Q, R, and P&Q&R.

35 Erin Kelly ("The Burdens of Collective Liability," in *Ethics and Foreign Intervention*, ed. D. K. Chatterjee and D. E. Scheid [Cambridge: Cambridge University Press, 2003]) suggests that many theorists emphasize corporate responsibility at a cost to individual enactor responsibility; I hope it is clear that my line is very different.

36 See Pettit, *Theory of Freedom*, chap. 5.

Philip Pettit

What is the company to do? Under the procedure described, it will take a majoritarian vote on each of the first three questions and then let the votes on those questions determine the issue about the pay sacrifice. The chair will take a show of hands on each of the first three questions, we may imagine, and then, identifying what they in logic require – I assume that people vote as they judge – announce that the corporate body has opted for making the pay sacrifice. The chair will not call for a vote on the issue of the pay sacrifice and, in ignorance of the fact that it is personally rejected by each of the members of the group, may even congratulate the group on its solidarity with the workers under threat and its public spirit in coming to their support.

But suppose now that some external parties have a complaint against the group, say, the spouses of the less-well-off workers, who think the pay sacrifice unfair. Whom, if anyone, can they hold responsible and blame for the line taken? Whom can they remonstrate with? Not the individuals in their personal right, since each can point out, the chair included, that he or she was actually against the pay sacrifice and that they were not in a position, as well they may not have been, to see the likely effect of the procedure they followed.³⁷

The spouses in this example can only blame the corporate group as a whole. The corporate organization has routines in place whereby the chair is triggered to represent the group in announcing the decision for a pay sacrifice and in putting that decision into effect. Since the corporate body is an agent that makes things tough for the less-well-off members, since it is in a position to recognize that reason as arguing against the line it takes, and since it is in reason-sensitive control of what it does, there can be no doubt about its responsibility. It is corporately responsible for what it did, even while the individual enactors may claim only a very diminished level of personal responsibility for the outcome.³⁸

VI. CONCLUSION: FROM CORPORATE TO COLLECTIVE RESPONSIBILITY

In conclusion it may be useful to consider briefly the issue of collective responsibility, for our discussion of corporate responsibility throws a certain light on that wider topic.³⁹ Although the discussion has been focused sharply on how incorporated groups can and should be held responsible for what they do, it suggests a reason

37 There is a problem of no hands, we might say, rather than the more familiar problem of many hands. See D. F. Thompson, "Moral Responsibility of Public Officials: The Problem of Many Hands," *American Political Science Review* 74 (1980): 905–16, and *Political Ethics and Public Office* (Cambridge, MA: Harvard University Press, 1987); M. Bovens, *The Quest for Responsibility: Accountability and Citizenship in Complex Organizations* (Cambridge: Cambridge University Press, 1998).

38 For a distinct but congenial lesson about corporate responsibility, to the effect that the greater powers of institutional agents may occasion more serious obligations in spheres such as that of global justice, see M. J. Green, "Institutional Responsibility for Global Problems," *Philosophical Topics* 30 (2002): 79–95.

39 For a sampling of the literature on this topic, see Larry May and Stacey Hoffman, eds., *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics* (Savage, MD: Rowman & Littlefield, 1991).

why apparently looser groupings can and should be held responsible in a similar way.

There are many cases of harms done by people in aggregate, where we have no inclination whatsoever to imagine that there is a group agent that we should hold responsible: an entity in respect of which we might feel resentment or indignation, as we contemplate the ill done. It may be, for example, that our species wiped out Neanderthal competitors about thirty thousand years ago. While we may feel regret that no members of that other species remain in existence today, it would be intuitively crazy to blame the human race, as if it were a group agent.

But there are cases where something close to holding a group responsible seems appropriate, even when the group does not constitute a group agent, or at least not a group agent that is cast in the ordinary mold. When bad is done – and that is the usual trigger – these are cases where we do want to speak of a collective guilt, distinct from the guilt of the individuals in the collection. There is something about the group that makes it appropriate to adopt the stance that we would normally reserve for individual and corporate agents: that is, for agents that clearly satisfy the three conditions for being fit to be held fully responsible.

The cases where this is so typically involve national peoples, as distinct from governments, or religious congregations, as distinct from their episcopacies or elders or imams. Where the states and episcopacies would count normally as corporate agents, that is not clearly so with peoples or with congregations, yet we do often want to find these groupings collectively responsible as well. We might want to blame Christians for the treatment of Jews in Western history or the treatment of native populations in the colonial countries where they sought to proselytize. We might want to blame New World colonists and peoples, for example in America or Australia, for the shameful treatment of indigenous populations or the German people for their acceptance of Nazi atrocities.

Is the ascription of group-level responsibility in such cases sensible? I think it may be. The reason is that the groupings involved here can be seen as potential or embryonic group agents, distinct from the group agents – the governments or episcopacies – that do the immediate harm. Those group agents pose as spokesbodies for the larger groupings, and to the extent that their claim to authority is not challenged, they have the tacit authorization of the members of the larger groupings. Thus, in each case the larger grouping may be seen as an embryonic group agent, working under a constitution that gives the spokesbody more or less unchecked and uninhibited status in determining the attitudes and actions of the whole.

But even if the larger grouping is an embryonic group agent, there is still a problem with holding it responsible. By hypothesis, everything decided and done by that grouping is decided and done, without check or inhibition, by the spokesbody. And so it may seem that, having ascribed enactor responsibility to that body, and perhaps to its active or passive collaborators among the people or congregation, there is no purpose served by going on to ascribe corporate responsibility to the grouping as a whole. The case will be unlike those examples rehearsed earlier in which there is a

shortage of enactor responsibility among members and so very good reason to focus on the corporate responsibility of the group as a whole. Here, by hypothesis, the enactor responsibility of members, in particular of the dictatorial spokesperson, leaves no shortfall in the responsibility stakes.

What reason can there be, then, for persisting in the ascription of a corporate form of guilt to a people or nation or to a body of believers? I think that doing so can have a developmental rationale, to return to a thought from the beginning of the article.⁴⁰ To refuse to ascribe collective responsibility to the grouping as a whole, on the grounds that the evil done was done entirely by the spokesperson, would be to miss the opportunity to put in place an incentive for members of the grouping as a whole to challenge what the spokesperson does, transforming the constitution under which they operate: making it into a constitution under which similar misdeeds should be less likely.⁴¹ By finding the grouping responsible, we make clear to members as a whole that unless they develop routines for keeping their government or episcopacy in check, then they will share in the corporate responsibility of the group; even if they have little or no enactor responsibility, they will have member responsibility for what was done. By finding the grouping responsible in such a case, indeed, we will make clear to the members of other groupings in the same category that they too are liable to be found guilty in parallel cases, should the body to which they belong bring about one or another ill.

This developmental rationale for ascribing group responsibility will be all the more powerful, of course, if the ascription of guilt is attended by a penal sanction of some kind. By way of parallel, think of the rationale for finding a commercial corporation responsible as a whole for some misdeed, rather than just finding the board or management responsible. Doing so is likely to provide an incentive for shareholders in that corporation, or in any similar corporation, to establish checks on the board and on management. It is likely, then, to elicit the sorts of capacities that will truly equip the group as a whole to be fit to be held responsible.⁴² What is true in this respect of the large commercial corporation is true equally of the citizens of a country and the believers in a church.

Once we recognize the developmental rationale that may make sense in these cases of the ascription of collective responsibility, we can envisage its extension to other cases where the collection that is held responsible falls well short of being even an embryonic group agent. Think of the school group whose members are told that they will all be held responsible if there is any sign of bullying in their midst or the loose professional association that is held accountable for the misbehavior of any member or the neighborhood that is held responsible in the public press when those

40 See D. P. Tollefsen, "Participant Reactive Attitudes and Collective Responsibility," *Philosophical Explorations* 6 (2003): 227.

41 Compare D. J. Levinson, "Collective Sanctions," *Stanford Law Review* 56 (2003): 345–428, for an argument about the beneficial effects of imposing sanctions on all the members of a collection "to motivate them to identify the guilty individuals in their midst" (348).

42 For an argument that corporations can be made more reactive to internal fault and shortfall in this fashion, see Brent Fisse and John Braithwaite, *Corporations, Crime, and Accountability* (Cambridge: Cambridge University Press, 1993).

who live there indulge in certain demonstrations, say of a racist character, or indeed the generation that is held responsible for the overuse and potential loss of antibiotics. It may not be strictly appropriate to hold such loose groupings responsible, since various of the conditions necessary for fitness to be held responsible will be missing. But holding it responsible may actually prompt the grouping to incorporate and organize against the condemned behavior.

We are naturally disposed to ascribe responsibility, it appears, not just to responsible agents and agencies but also to ‘responsibilizable’ entities – not just to agents that are fit to be held responsible but also to entities that are capable of being made fit to be held responsible. Perhaps grounded in the role that it plays in scaffolding the development of children, this disposition may serve us equally well in prompting social groupings to assume a corporate or quasi-corporate form.