III—NON-DISTRIBUTIVE BLAMEWORTHINESS

THOMAS H. SMITH

I adapt an old example of Frank Jackson's, in order to show that it is not only possible that actions with different individual agents are sub-optimal when each is not, but that they are impermissible when each is not, and blameworthy when each is not.

I

Introduction. The sort of distribution I have in mind was originally conceived of as a property of linguistic items. One of these was said to be *distributive* whenever it applied to each of the items to which it applied; this was once thought to be a feature of certain terms:

> A term is said to be distributed if reference is made to all the individuals denoted by it … (Keynes 1906, p. 5)

These days it is more commonly conceived of as a feature of predicates:

> Let P be any plural term … the predicate F is *distributive* if and only if F(P) is equivalent to 'Each one of P is F' … (Oliver and Smiley 2001, p. 289)

I reconstrue distribution as a property of property possession (where properties are conceived of as *universals*). A property is distributively possessed by some things just if it is possessed by each of them. Hence, things are distributively F just if they are each F. A property is non-distributively possessed by some things if it is possessed by them and not by each: things are *weakly* non-distributively F just if they are F and at least one of them is not F, and *strongly* non-distributively F just if they are F and each one of them is not F. Throughout this paper, I will have this stronger notion in mind.

Many properties are on occasion (strongly) non-distributively had: my books are heavy but each is not. These include psychologi-

cal properties: we all know of committee members who decide to do something (hire a candidate, say), even whilst no one of them thus decides, not least because no one of them has the authority so to do (each may *concur* with the decision, but no one makes one).[1] Evaluative properties too are, on occasion, non-distributively had: the points in Seurat's *pointillist* painting *Un dimanche après-midi à l'Île de la Grande Jatte* are beautiful but each unbeautiful.

On a natural understanding of these examples, the properties of heaviness, decision-making and beauty are collectively as well as non-distributively had: the books are collectively heavy, the committee members collectively decided, the points collectively beautiful. But non-distribution is neither necessary nor sufficient for collectivity. To see that it is not necessary, observe that Lennon and McCartney collectively wrote songs, but did not non-distributively write songs, as they wrote songs separately as well as together. To see that it is not sufficient, observe that Rodgers, Hammerstein, Lerner and Loewe non-distributively wrote musicals, but did not collectively write musicals; rather, Rodgers and Hammerstein collectively wrote musicals, and Lerner and Loewe collectively wrote musicals. As I see it, a property is *collectively* possessed by some things just if there is at least one instance of it that is possessed by (all of) them. There is at least one instance of heaviness that is my books' heaviness, one instance of beauty that is Seurat's points' beauty, one instance of decision-making that is the committee members' decision-making, and one instance of songwriting that is Lennon and McCartney's (even though there are also instances that are Lennon's and instances that are McCartney's); by contrast, there is no instance of musical writing that is Rodgers, Hammerstein, Lerner and Loewe's.[2]

Actions are sometimes, if not exactly beautiful like Seurat's points, pleasing or displeasing (repugnant, sickening, etc.) to our *moral* sensibilities. And we can all think of cases where some actions—with different individual agents—are *non-distributively* thus displeasing: sometimes, when many nods or lapses, each of trifling significance, together constitute or cause a catastrophe (a rail disaster, the election of a nasty party, an uninhabitable planet), they will offend and displease us even though each does not.

What is perhaps less clear is that any such plurality of actions is

---

[1] Here I follow Velleman (1997, pp. 29–30).

[2] I expand on these remarks in Smith 2009*a*.

non-distributively *blameworthy*, i.e. that the actions are blameworthy even though each is not. As I understand it, any individual agent's action is blameworthy just if it is appropriate to blame him for performing it. More generally, for any actions $A_1, ..., A_n$ the property of blameworthiness is possessed by $A_1, ..., A_n$ just if it is appropriate to blame whoever performed them for performing them:[3] it is *distributively* possessed by $A_1, ..., A_n$ just if, for each of $A_1, ..., A_n$, it is appropriate to blame whoever performed it for performing it, and it is *non-distributively* possessed by $A_1, ..., A_n$ just if it is appropriate to blame whoever performed them for performing them, even though, for each of $A_1, ..., A_n$, it is not appropriate to blame whoever performed it for performing it.

When one reflects coolly on rail disasters, elections of nasty parties and the like, one generally finds oneself inclined to adopt one of two responses, namely (i) that whilst, say, *dismay*, *lamentation* and *distaste* may be appropriate reactions, *blame* is not appropriate, and (ii) that it is appropriate to lay *some* blame on *some* individual agent, for his (perhaps small) part in the catastrophe. No case of non-distributive blameworthiness (henceforth 'NDB') would legitimate either response: a genuine case would be one in which it is appropriate to blame the agents for their actions and yet not appropriate to blame any one of them. I will present a possible case of NDB.

II

Jackson's Case. Imagine a world *w* much like our own, but in which cars can only be directed to go at either 80mph or 60mph. In this world, some agents $a_1, ..., a_n$ are driving to work on the motorway. Stopping is not an option, and there are no turn-offs. Each driver drives at 80, which is dangerously fast: there is a 1 in 10,000 chance of a fatal accident if all drive at 80. There is only a 1 in 100,000 chance if all drive at 60. The risk is greatest, however, if some drive at 80 and some at 60, because of all of the overtaking that would then ensue: the chance of a fatal accident in this circumstance would be at

---

[3] It is not sufficient for the blameworthiness of $A_1, ..., A_n$ that it is appropriate to blame whoever performed them *for them*. If you and I coerce each other into, say, acts of deception, these acts are non-blameworthy, on account of their being coerced; but it is appropriate to blame you and me for them.

least 1 in 5,000 (the exact figure would depend on the ratio of fast to slow drivers, but as this approximates to 1:1, the speed of any one driver makes no significant difference).

Imagine now that, in w:

> (J)   For each one of $a_1, ..., a_n$, were he to drive at 60, the rest of $a_1, ..., a_n$ would still drive at 80.

What is imagined, then, is a world containing some agents $a_1, ..., a_n$ such that:

> (J_1)   Were $a_1, ..., a_n$ to drive at 60, the road would be safer.
>
> (J_2)   Were a_1 to drive at 60, the road would be less safe.
>
> ...
>
> (J_{n+1}) Were a_n to drive at 60, the road would be less safe.

'J' is for Frank Jackson. He imagined a world like this over twenty years ago:

> There is a steady stream of traffic going to work. Everyone is driving at 80 ... It would be safer if everyone was driving at 60 ... But [for] each person ... if he or she were to drive at 60, everyone else would still drive at 80, and so a lot of dangerous overtaking would result. (Jackson 1987, p. 102)[4]

It seems to me that Jackson imagined a *possible* world. Of course, the antecedent of (J_1) is equivalent to the conjunction of the antecedents of (J_2), ..., (J_{n+1}), and the consequent of (J_1) contradicts the consequent of each of (J_2), ..., (J_{n+1}). But on standard views of the logic of counterfactuals, claims of this sort are, even if read non-vacuously (i.e. as having possible antecedents), formally consistent.[5]

I also see no reason to doubt that Jackson's world is one in which (J_1), (J_2), ..., (J_{n+1}) are non-vacuously true, i.e. from its standpoint, each has a possible antecedent.

Jackson does not say why his drivers drive at 80, but it is easy to imagine why. Indeed, we can—and I hereby do—coherently attribute to each of $a_1, ..., a_n$ perfect rationality, an overriding concern

---

[4] See also Jackson (1988, pp. 266–7). I am heavily indebted to these papers. Similar cases are discussed in Regan (1980), Feldman (1986), Parfit (1988), and Bacharach (2006).

[5] See Lewis (1973, §§ 1.2, 1.4, 1.8).

with road safety, beliefs which mirror, and evidence which indicates, all and only the relevant facts, *and* a motive for driving at 80, and for continuing to drive at 80 were another one of them to slow down to 60. We need only suppose that, even though each driver recognizes the safety benefits of uniform slow driving, each believes, as it happens correctly, that everyone else will drive fast, perhaps because each believes, as it happens correctly, that everyone else believes that everyone else will drive fast, perhaps because each believes, as it happens correctly, that everyone else believes that everyone else believes that everyone will drive fast, etc. These beliefs are collectively self-fulfilling: each drives at 80 because he believes that the others will do so and he reasonably prefers to drive at the same speed as them. And if one of $a_1, \ldots, a_n$ were to slow down to 60, the others would, quite reasonably, minimally modify their beliefs and preferences so they concern what is believed and done, not by everybody else, but by *most* everybody else: they would come to believe, truly, that most everybody else will drive fast, and prefer, reasonably, to do the same.

Someone might dispute the coherence of my attributions on the following grounds: were a driver to slow down to 60, each of the others would, if rational, see this as a salient 'focal point' for coordination (see Schelling 1960), which gives him most reason to slow down too, as it entitles him to expect that others will also see the driver's act as a focal point for coordination, and hence to expect that they too will slow down, given their common overriding concern with road safety and (among their many relevant true beliefs) their common recognition that their all slowing down is their safest option. Furthermore, *since* this is so, and since the drivers, trapped as they are in discrete, moving vehicles, are deprived of conventional means of communication, each driver has most reason to create, at his earliest opportunity, a focal point of the sort described, by slowing down to 60, thereby implicitly inviting the others to do likewise. Surely then I incoherently attribute to each driver perfect rationality, an overriding concern with road safety, relevant true beliefs, etc., *together with* a motive for driving at 80 and for continuing to drive at 80 were another driver to drive at 60.

I answer that, whilst it may be rational to attempt to create a focal point by slowing down to 60, and rational, should another slow down to 60, to interpret this as a focal point, it is just as rational not to do either of these things. It is just as rational to believe that everyone else will drive at 80, and to prefer to do the same, and, should

someone else slow down to 60, to minimally modify one's beliefs and preferences and believe that everybody else, excepting him, will drive at 80, and prefer to do the same.

Jackson's case provides me with the raw materials with which to construct my case of NDB. But it is worth attending to what Jackson himself says about his case:

> For each individual the right action is to keep driving at 80, so avoid dangerously disrupting the traffic flow; yet … we have a wrong group action—everyone together driving at 80 …. (Jackson 1987, pp. 102–3)

This is a moral intuition with metaphysical presuppositions, as, for Jackson, a 'group action' is a single complex action, composed of a plurality of actions,[6] and he seems also to hold that a group action has a single complex agent, a 'group', which is composed of a plurality of agents.[7] Whilst it would be hard to deny that Jackson has successfully imagined a possible situation featuring a *plurality* of actions of fast driving, each with its own agent, one might reasonably doubt that he has imagined a possible situation in which some such plurality of actions compose a complex action, with a single complex 'group' agent, composed of a plurality of agents.

Charitably purged of its dubious metaphysics, Jackson's claim is:

> For each individual the right action is to keep driving at 80, so avoid dangerously disrupting the traffic flow; yet we have a wrong plurality of actions—a_1's driving at 80, …, a_n's driving at 80.

In other words, wrongness is non-distributively possessed by the acts of driving.

This is a somewhat indeterminate claim, as 'right' and 'wrong' are used by philosophers and folk to identify several distinct notions. I will proceed by defending, in the next three sections, the following three precisifications of Jackson's intuition:

> For each individual an *optimal* action is to keep driving at 80, yet we have a *sub-optimal* plurality of actions.

---

[6] Indeed, he holds that 'any old aggregation or mereological sum of individual actions counts as a group action' (Jackson 1987, p. 93).

[7] Jackson (1988, p. 267): 'As groups are to their constituent individuals … so individuals are to their temporal parts.'

For each individual a *permissible* action is to keep driving at 80, yet we have an *impermissible* plurality of actions.

For each individual a *non-blameworthy* action is to keep driving at 80, yet we have a *blameworthy* plurality of actions.

I proceed thus partly because my defence of the third precisification rests, in part, on elements of my defences of the other two, but also because I hope that my discussion brings differences between the relevant moral properties into clearer focus.

My defences will derive support from a series of stipulations that add to the description of w so far given. If these stipulations about w, together with what has already been said about w, describe no possible world, this will of course undermine my project. I try to anticipate and protect against potential objections to this effect, often by making further stipulations. The effect is to transform Jackson's case from one approximating a real-world case into one remote from actuality. I confess to some disquiet about this; sadly, I cannot here defend the empirical claim that NDB is actual.

III

Non-Distributive Sub-Optimality (NDSO). There is a measure of action, namely that of magnitude of contribution to road safety, relative to which a_1, ..., a_n's acts of driving are evaluable as *non-distributively sub-optimal*. For these acts make a contribution to road safety that is sub-optimal, but each of them makes a contribution to road safety that is optimal and hence not sub-optimal.

Let me try to render that thought more precise. On what I take to be a reasonably intuitive notion of an *option*, for any agents x_1, ..., x_n (in any world y) an *option* for x_1, ..., x_n is some compossible token action(s), each with one or more of x_1, ..., x_n as its agent(s), such that each of x_1, ..., x_n is the agent, or among the agents, of at least one of them. The action(s) of which an option consists can be, in part or in whole, actual or non-actual.[8] I say that a_1, ..., a_n's acts of driving make a sub-optimal contribution to road safety, because, by (J_1), there is a safer option for a_1, ..., a_n than what they do, namely to all

---

[8] Strictly speaking, I should speak of options for agents *at times*. I shall ignore this complication.

drive at 60. And I say that each act of driving makes an optimal con-
tribution to road safety because, for each one of $a_1, ..., a_n$, by $(J_2), ...,$
(J_{n+1}), there is no safer option for him than what he does. (Recall that
each driver has but two options, to drive at 80 and to drive at 60.)
The background assumption here is that when we compare contribu-
tions to road safety made by agents' options, we compare—from the
standpoint of those agents' world—what *would* happen *were* they to
pursue those options (or what *does* happen, in the case where an op-
tion *is* pursued).

Now, let $a_1, ..., a_n$'s journeys be non-urgent, let them be bound by
no promises, let them derive no morally significant pleasure from
speed, etc.; in short, let magnitude of contribution to road safety be
the only morally significant measure of their driving in *w*. Given this
stipulation, the best option for $a_1, ..., a_n$ is to all drive at 60. So their
actions are non-distributively sub-optimal *simpliciter*: there is a bet-
ter option for them than what they do, yet, for each of them, no better
option for him than what he does.[9]

I do not know how interesting a result this is, but I pause to note
that we have here a case of NDSO that makes simpler assumptions
about optimality measures than certain other more familiar sorts of
case that may also be argued to be cases of NDSO.

Compare the Prisoner's Dilemma, which we may represent thus
(where the numbers on the left represent my valuings, and the ones on
the right your valuings):

	You talk	*You stay silent*
I talk	−1, −1	2, −2
I stay silent	−2, 2	1, 1

If we both talk, then, by the measure of our own valuations, our ac-
tions are sub-optimal, but each is optimal. For each values 'both
stay silent' more highly than 'both talk', and each values his own
talking more highly than his own silence, in that, regardless of
whether the other talks or stays silent, each values the other's doing
this together with his own talking over the other's doing this togeth-
er with his own silence. Notice, however, that there are *two* meas-
ures in this case: one supplied by my valuations, and one supplied

---

[9] See Jackson (1988, p. 266): '… it would be better if everyone drove slower … for each per-
son it would be worse if he or she drove more slowly'.

by yours. (This fact is obscured by my use of the phrase 'the measure of our own valuations'.) Each agent's talking is optimal relative only to his *own* measure. By contrast, in Jackson's case, each driver's driving is optimal, and the drivers' driving is sub-optimal, relative to a single measure, that of magnitude of contribution to road safety (which, we are supposing, is the only relevant measure). Hence, the drivers' NDSO is *impartial* in a way in which the prisoners' NDSO is not.

Another sort of candidate case of NDSO employs a single measure, but assumes that marginal differences by that measure are insignificant. This might be the right way to think about hypothetical cases in which each of very many agents free-rides on contributing to some incremental public good. Suppose that each of us defaults on his taxes, and that the only morally significant measure of these actions is that of magnitude of contribution to public funds. The defaulting acts are plainly sub-optimal, by this measure. But it may be argued that each is optimal—not *best* perhaps, but *unbettered*—as, no matter how many pay, any one act of payment makes a negligible contribution to public funds. There is much to be said about such cases. All I want to note here is that Jackson's case employs a measure of action by which, for each driver, there is a *non-negligible* difference between his contribution to road safety were he to slow down, and his contribution to road safety were he not to: were he to slow down, the chance of a fatality would be at least 1 in 5,000. Were he to drive at 80, it would be 1 in 10,000. This is to the credit of Jackson's case, as marginal differences notoriously raise many philosophical puzzles, which are far from solved.

In the Prisoner's Dilemma and free-riding cases, each of some sub-optimal actions is (arguably) optimal in this sense: when we compare its agent's options *whilst making no assumptions about what the others do or would do*, we see that, by the relevant measure, he pursues his best option, or at least an unbettered option. In Jackson's case, each of some sub-optimal actions is optimal in this sense: when we compare its agent's options *whilst taking into account, for his actual option, what the others do, and for each non-actual option, what the others would do were he to pursue it*, we see that, by the relevant measure, he pursues his best option. Perhaps there are cases regarding which the former notion of optimality is most appropriate. It seems to me, however, that the latter notion is intuitively applied to Jackson's case.

IV

Non-Distributive Impermissibility (NDI). Agents act impermissibly just in case what they do is obligation-violating.[10] And I take it that an obligation is a certain sort of *demand*: not a demand of the sort that is made of agents by agents, but a demand of the sort that is made of agents by the world.[11] Less picturesquely, an obligation is a demand of the sort that is made of some agent(s), by some fact or state of affairs. In some cases, the fact in question will be that some agent(s) have made a like demand. But not everything demanded of us by some fact—not everything that we are obligated to be or do— is demanded of us by others. And not everything demanded of us by others is demanded of us by any fact.

This talk of facts issuing demands may strike readers as fanciful. But reasons realists should not balk at it. Reasons realists believe that facts supply agents with (more or less powerful) reasons for beliefs, actions, and so on, and many suppose that some such facts supply agents with reasons that are conclusive or binding.[12] It is not fancy talk to say that facts of the former sort *recommend* beliefs, actions and so on to agents, and that facts of the latter sort *demand* beliefs, actions and so on, of agents.[13]

Or rather, if this is fancy talk, there is a point to it. It is illustrative fancy talk, for the reasons that I characterize as recommendations and demands tend to mimic or (depending on your metaphysic) be mimicked by *reasonable* actual and hypothetical recommendations and demands emanating from other agents. The principle I am fumbling for is, at first blush, that something is recommended to, or demanded of us ('by the world') *if* someone could reasonably recommend it to, or demand it of, us.

No doubt that first blush principle requires modification. For one thing, it may be that someone in possession of false beliefs, or misleading evidence, could reasonably recommend or demand things

---

[10] This may be an oversimplification. Perhaps agents act impermissibly relative to the facts just in case what they do is obligation-violating, impermissibly relative to their beliefs just in case they believe that what they do is obligation-violating, and impermissibly relative to their evidence just in case their evidence indicates that what they do is obligation-violating. I can ignore this complication because, as I have said, $a_1, ..., a_n$ have beliefs that mirror, and evidence that indicates, all and only the relevant facts.

[11] Here I follow Pink (2004).

[12] See, for example, Scanlon (1998), Raz (1999), Dancy (2000), Parfit (forthcoming).

[13] I owe this use of the recommendation–demand contrast to Pink (2004).

that are not recommended or demanded 'by the world'. Also, it may be that some things that we have reason to be or do could not be recommended to or demanded of us, without *unreasonable* busy-bodying, by one with no *interest* in our being or doing such things. It may also be that, where we share a history of reciprocity or betrayal with another, or bear a special relation of friendship, love, authority, or some such to them, the mere possibility of their reasonably recommending or demanding something to or of us does not entail that this *is* recommended to or demanded of us 'by the world' (although, perhaps, in some sorts of case, it follows that it *would* be recommended or demanded, *were* the person to make the relevant recommendation or demand). There are also nice questions as to how much of the actual world to 'hold fixed' when we hypothesize a reasonable recommendation or demand. So the principle for which I am fumbling is hard to state more determinately than this: if a certain sort of person—*one whose beliefs mirror, and whose evidence indicates, the relevant facts, who has an interest in the matter, and who shares no history with us, and bears no special tie to us*—could—*in a world much like ours*—reasonably recommend or demand that we be some way or do some thing, then some fact recommends, or demands, that we be that way or do that thing.

The defence of the principle is simply this: *reasonable* recommendations or demands correspond with *reasons* that have the force of recommendation or demand.

It would be a cheek not to note the similarity between the connection here assumed between certain sorts of reasons on the one hand, and certain sorts of actual and hypothetical reasonable acts on the other, and the one claimed by Scanlon's celebrated principle that 'an act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behaviour that no one could reasonably reject …' (Scanlon 1998, p. 153). I think that by 'any set of principles' Scanlon means *some set of principles*, not *every set of principles*.[14] However that may be, on the assumption that to reasonably reject principles entailing a prohibition one must reasonably reject that prohibition, Scanlon's claim entails this one: an act is wrong if no one could reasonably reject its prohibition; equivalently (it would seem) an act is wrong if someone could reasonably reject its permission. This claim is very

[14] Parfit (2003) reads him thus.

similar to my assumption that an act violates an obligation if it is contrary to a demand that a person—of a certain sort—could reasonably make. It is perhaps worth noting, then, that, unlike Scanlon, I do not suppose that the possibility of certain sorts of reasonable attempt to preclude *explains* normative or deontic phenomena or facts. Scanlon thinks that his principle helps to explain something about the acts that are, by that principle, wrong: what their wrongness consists in, or the significance of their being wrong in the way that they are.[15] I make no like claim. I take no stand as to whether facts about the world's recommendations and demands are explained by facts about what others can reasonably recommend to us, and demand of us, or whether the direction of explanation runs in the other direction, or in both directions, or in neither.

Now, Jackson's hypothetical world w becomes no less possible if we add to its description the existence of someone—call him 'b'—whose beliefs mirror, and whose evidence indicates, the facts about $a_1, ..., a_n$'s driving, who has an interest in their safe driving (children who live near the road, let's say), who shares no history with them, and bears no special tie to them, and who demands, of them, that they all drive at 60, using—we may suppose—a PA system hooked up to loudspeakers (one in each car).

In introducing my principle about reasonable recommendations and demands, I spoke only of its application to the actual world. Perhaps the principle fails in some worlds. But w is, apart from the activities of $a_1, ..., a_n$ and b, as similar to the actual world as you like. So there is no reason to doubt that the principle applies to w. And, from the standpoint of w, it is not just that someone of the requisite sort *could* demand of $a_1, ..., a_n$ that they all drive at 60, someone, b, does demand this. Our question must then be whether this demand is reasonable. It seems to me that given b's interest in the matter, the fact that there is a 1 in 10,000 chance of a fatal accident if all drive at 80, and a 1 in 100,000 chance if all drive at 60, and the fact that no driver is in a hurry, or is deriving any morally significant pleasure from driving fast etc., b's demand is more than reasonable. I infer that $a_1, ..., a_n$'s driving in w is impermissible: the world demands of $a_1, ..., a_n$ that they drive otherwise, and they violate this demand.

I anticipate an objection: b's demand is unreasonable because, as

---

[15] He makes such claims, albeit in slightly different ways, in Scanlon (1998, Introduction; 2003).

the drivers are not in a position to communicate, they cannot coordinate their response. I respond by stipulating that the PA is also an intercom system, and that it is common knowledge among a_1, ..., a_n that b manipulates it thus: following his demand there are ten minutes, during which the drivers can communicate, but cannot drive at any speed other than 80, after which b reissues his demand, communication ceases, and the drivers' options are as before, drive at 60 or drive at 80. It strikes me as obvious that b could reasonably use such a system to demand of the drivers that they all drive at 60.

Now, a_1, ..., a_n's acts of fast driving are non-distributively impermissible just if they are impermissible and each is permissible. I have argued that they are impermissible. It also strikes me as overwhelmingly plausible that each is permissible. For each driver has only one other option, which is to drive at 60, and no one could reasonably demand of a driver that he do *this*, for, as each driver believes, and as his evidence indicates, were he to slow down, the road would then be less safe, and there would be no compensating benefits of any significance in the circumstances. Let me be clear about what I am claiming: it is not that no driver could reasonably slow down, but that no one could reasonably demand of any one driver that he slow down. And given this, I see no reason to think that the *world* demands this of any one driver.

The introduction of a discussion period does not alter this fact, provided that, as I now stipulate, the drivers make no commitments to drive at 60 during that time.

The stipulation may provoke a fresh objection, which is that each driver is obligated to use the discussion period to publicly commit himself to slow driving, by promising or announcing an intention to slow down, and that, if he does not do this, his subsequent fast driving *inherits* some of the impermissibility of his failure to thus commit.[16] I do not accept that any driver is thus obligated. Perhaps it would be rational and reasonable to make a public commitment of the sort envisaged, in the hope that others will view this as a focal point for coordination, make like commitments, and go on to honour them.[17] But no one could reasonably *demand* of any driver that he thus commit himself, as, by (J), were any driver to slow down,

---

[16] Chloë Fitzgerald and David Wiggins pressed similar objections.

[17] Although the 'final stipulation' that I am to make in the next section, on pp. 52–3, renders this doubtful.

this would precipitate the worst case scenario, in which drivers drive at variegated speeds.

Still, one might think that each driver is obligated to at least *conditionally* commit himself to drive at 60 (the condition being that everyone else makes a like conditional commitment), or to *conditionally* partake of an agreement or collective decision to all drive at 60 (the condition being that everyone else likewise conditionally partakes).[18] By so doing, he would ensure that he is committed to driving more slowly *when and only when* everyone else is likewise committed, and so—it would seem—would minimize the risk of the worst case scenario (drivers drive at variegated speeds), whilst raising the chances of the best case (uniform slow driving).

But neither conditional commitment to drive at 60, nor attempted participation in an agreement or collective decision to drive at 60, are reasonably demanded of any driver, *if*—as I shall now stipulate—as every driver believes, and as their evidence indicates, were the drivers to thus commit to all drive at 60, none would believe that the others will drive at 60. More particularly, my stipulation is that (as each believes, and as their evidence indicates) each would have *no belief* (of any degree of certainty) and *no evidence* as to how others will drive, and, furthermore, that each would respond to such uncertainty by choosing his speed by some random procedure akin to coin-flipping, for each would take himself have no more reason to drive at one speed rather than the other, his perfect rationality, concern with road safety, true beliefs and non-misleading evidence notwithstanding. Each would take himself to have most reason to drive at 60 *if* the others do, and most reason to drive at 80 *if* the others do, but no reason to adopt either view of how others will drive. For each would believe, truly, that every other driver has no belief and no evidence as to how others will drive, and so takes himself to have no more reason to drive at one speed rather than the other, perhaps because each would believe that every other driver *believes* that every other driver has no belief and no evidence as to how others will drive, and so takes himself to have no more reason to drive at one speed rather than the other, and so on.

---

[18] I mean to here leave open whether these two sorts of conditional commitment are one and the same. My understanding of conditional commitment, agreement and collective decision owes much to Gilbert (see esp. Gilbert 1996, chs. 6, 12, 13) and to discussion of her views by Velleman (1997), Roth (2004), and Black (2004; 2007). For my own views on these phenomena, see my (2009*b*).

No doubt, for each driver, there would be a temptation to reason as follows: for each driver other than me, his driving at 60 and his driving at 80 are equiprobable; so there is a small chance that everyone else drives at 60, and an equally small chance that everyone else drives at 80; so whilst, whatever I do, it is probable that this will lead to the worst case scenario in which we drive at variegated speeds, if I drive at 60, there is a small chance that this will lead to the best case scenario in which we fulfil our obligation to all drive at 60, whereas, if I drive at 80, there is an equally small chance that this will lead to the second best case scenario in which we all drive at 80, and no chance at all that it will lead to the best case scenario; therefore, I will drive at 60. But no driver *would* reason thus. For, were he to, then—given his *ex hypothesi* perfect rationality, and belief in, and evidence for, the relevant fact that every other driver is perfectly rational and has an overriding concern with road safety—he would predict that every *other* driver performs the same reasoning, and hence that everyone is certain to drive at 60. But this prediction would contradict a premiss of the reasoning, namely that, for every other driver, his driving at 80 and his driving at 60 are equiprobable. And a perfectly rational driver would not accept contradictory claims.[19]

To clarify, I am not amending my stipulation that each driver believes that the others will drive at 80. I am adding the stipulation that *were the drivers to commit to all drive at* 60, each would have no belief and no evidence as to how others will drive.

Might the drivers' committing to all drive at 60 not *itself* be evidence as to how they will drive? No, by stipulation, each driver would just *know*, perhaps by induction, or by some magical means, that their commitment has no evidential worth.

Might any driver's fast driving inherit impermissibility from his disposition to take himself to have no more reason to drive at one speed rather than the other, were the drivers to commit to all drive at 60? After all, one might think that, were the drivers thus committed, neutrality between options would impermissibly dishonour the commitment(s) made. I doubt that any driver's neutrality would be impermissible in the circumstances, as (a) for each driver, his neutrality is a reasonable response to the neutrality of the others, (b) these are highly likely to be circumstances in which the drivers drive at varie-

---

[19] I am indebted to Bacharach's discussion (2006, p. 52) of a similar response to the 'Hi-Lo paradox'.

gated speeds, and hence ones in which, for each driver (i) others have defaulted on a commitment to all drive at 60, plausibly releasing him from any obligation not to likewise default, and (ii) his speed will not significantly affect the safety of the road. But even if I am wrong about this and neutrality would be impermissible, it does not obviously follow that each driver's disposition to be thus neutral is impermissible, and it is, I think, highly doubtful that any driver's *actual and otherwise permissible* action of driving at 80 inherits any impermissibility from a disposition of his, *which never manifests itself*, to be neutral in certain circumstances.

My strategy has been to argue that driving at 60 is not demanded of any one of $a_1, \ldots, a_n$, as each has to operate in an imperfect world in which he must coordinate his actions with those of others who have been stipulated to have beliefs, preferences, and dispositions to believe, prefer and act in certain circumstances (as when commitments are made), which render unreasonable demands made of him to (i) drive at 60, (ii) commit conditionally or unconditionally so to drive, and (iii) conditionally partake of an agreement or decision so to drive. But does not the necessity of making the best of an imperfect world also render unreasonable any demand made of the *drivers* to all drive at 60, contrary to what I have argued? No. For when we contemplate demanding something of some agent(s), facts about the psychologies of some *other* agent(s) are reasonably conceived of as part of the 'fixed background' against which the agent(s) we contemplate must operate. So, when we contemplate demanding something of any *one* of $a_1, \ldots, a_n$, facts about the psychologies of the rest of $a_1, \ldots, a_n$ are reasonably conceived of as part of the 'fixed background' against which that agent must operate. But when we contemplate demanding something of $a_1, \ldots, a_n$, facts about the psychologies of $a_1, \ldots, a_n$—and, in particular, facts about their 'judgement-sensitive attitudes' (Scanlon 1998, ch. 1) of belief, preference, etc., and their dispositions to hold such attitudes—are, however cranky those psychologies may be, not reasonably conceived of as part of any such background.[20] Hence such facts can render unreasonable demands made of any one of $a_1, \ldots, a_n$, without rendering unreasonable like demands made of $a_1, \ldots, a_n$.

---

[20] See Parfit (1989, pp. 510–11, n.41): 'When we ask whether each acted rightly, the circumstances include what others did. When we ask whether we acted rightly, this is not so.' See also Humberstone (1983; 1991).

I anticipate a new objection: someone might insist that, contrary to what has been argued, it just *follows* from the proposition that actions of driving at 60 are demanded of the drivers, that each of these actions is demanded of its designated driver. But this is not so. The most that can be said to follow is that each of the actions of driving at 60 is *one* of some actions that are demanded of the drivers, and we need not assume that to be one of some demanded actions is to be a demanded action. After all, to be one of some heavy books is not to be a heavy book. It may be granted that *if*, for the world to demand a plurality of actions, it must make a plurality of demands, each a demand for one action, *then*, if a plurality of actions are demanded, each one of them is. But I see no reason to accept that the world can demand a plurality of actions only by making a plurality of demands. An analogy may help: a school bully decides to demand of the class weaklings that they give him their lunch money; he might write several threatening notes, and give each weakling one of them, or he might write one note, and have it passed around the classroom. This *is* an analogy: it shows that one can demand a plurality of actions by writing one note, rather than a plurality of notes; I do not claim that there is a one-to-one correspondence between the number of notes issued by the bully and the number of demands that he makes (still less that there is there is always a one-to-one correspondence between the number of demands made by someone or something, and the number of material means—notes, speech acts, etc.—by which they are made). The point of the analogy is to show that, absent an account of how demands differ relevantly from demanding notes, we are entitled to doubt that if a plurality of actions is demanded, then a plurality of demands are made.

Jackson does not directly apply the concept of demand to his case. He does say that, whilst no driver ought to slow down, the drivers ought to slow down (Jackson 1988, pp. 266–7), and the context suggests that he understands this to be equivalent to the claim that, for the drivers, slow driving is obligatory, whereas for each driver, it is not obligatory. But these claims are not equivalent: things that ought to be done by things are not thereby obligatory. Suppose we replaced each of $a_1, ..., a_n$ with a robot. Plausibly, were these robots functioning well relative to a human purpose serving a human need (namely, the need for a safe road), they would slow down; in *that* sense slowing down is something that ought to be

done by them (rather as detecting viruses is something that ought to be done by anti-virus software). But—I take it—robots, like brutes, never violate *obligations*. They are never contrary to what the world demands of them, because the world demands nothing of them. (Of course, *we* sometimes make demands of robots, brutes, and the like, but any such demand is unreasonably made, and so is no evidence of anything being demanded by the *world*). If robotic counterparts of a_1, …, a_n can fail to do what they ought to do, without violating any demand, then establishing that a_1, …, a_n fail to do what they ought to do does not establish that they violate a demand.

Plausibly, nothing is demanded of robots, brutes, and the like because, whilst a robot or brute can act *in accordance* with putative demands (in a weak sense of 'act' that does not presuppose rational agency), it is unable to *heed* any demand. For to heed a demand is to exercise practical reason; it is to *acknowledge* the demand (perhaps after *deliberating* between options) and to *resolve* to act in accordance with it, before attempting so to do. And, I assume, no robot or brute has practical reason, or is capable of performing mental acts of acknowledgement, deliberation or resolution.

These reflections suggest a principle: for any possible thing(s), they are subject to a demand only if they could *heed* it. As a demand is a reason with conclusive or binding force, and assuming that heeding a reason is, or involves, motivation by it, this is a sort of 'internalist' principle. But its internalism is modest: the condition claimed to be necessary for a reason making a demand of some thing(s) is not that they *would* be motivated by it (were they to deliberate from certain motivations), but that they *could* be. The principle may strike one as a good (perhaps the best) explanation of why nothing is demanded of robots, brutes, and the like.

I suspect that the principle is too strong. For (*pace* 'ought entails can'), we can imagine a person who is so constitutionally selfish that there is *no* possibility of their heeding demands to help others, but who is, for all that, subject to such demands. More plausible is a weaker principle, namely, that some thing(s) are subject to a demand only if they have a capacity to heed demands in general. This would suffice to explain why nothing is demanded of robots or brutes. But I am content to grant the stronger principle for argument's sake. For what I want to point out is that this principle is no threat to my argument for the possibility of NDI. Agents a_1, …, a_n are more than capable of acknowledging (perhaps after deliberating

between options) that driving at 60 is demanded of them, and of re-solving and subsequently attempting to act in accordance with this demand, *simply by each doing so*. For an agent can deliberate, ac-knowledge, and resolve and attempt to act in accordance with de-mands that are not made of *him* alone. If a demand is made of you, I can deliberate, acknowledge, and resolve and attempt to act in ac-cordance with it. For example, if you have promised that you and I will come to Ian's party, then the fact of that promise demands, of you, that you and I come to Ian's party, and I can deliberate, ac-knowledge, and resolve and attempt to act in accordance with this demand. Likewise, agents can each deliberate, acknowledge, and re-solve and attempt to act in accordance with demands that are non-distributively made of them, and, by so doing, *heed* these demands (by each doing so). By contrast, were, *per impossibile*, demands made of a robot or brute, whilst a human agent might heed them, his acting thus would not suffice for the *robot or brute's* heeding any-thing, as individually performed acts of heeding can suffice for a plurality's heeding a demand.

An even stronger internalist principle would have it that for any possible *plurality* of things, they are subject to a demand only if they could *non-distributively* heed it. I doubt if this principle is true, but, if it is, I suspect that even it is no threat. For, as we have seen, decid-ing to do something can be non-distributive, and, plausibly, to de-cide to do something *because* it is demanded *just is* to end deliberation, acknowledge a demand, and resolve to act in accordance with it; all in all, it is to *heed* the demand in question. So perhaps heeding can be non-distributive.[21]

Our drivers violate an obligation, then, because whilst they *could* have heeded the demand that they all drive at 60, they either failed to acknowledge this demand, or acknowledged it but did not follow through by each thus driving. No driver violates an obligation, however, as no driver acts contrary to any demand that is made of him. Hence, I submit, we have in *w* a possible case of non-distribu-tive impermissibility.

---

[21] I recognize that these remarks are insufficient. I will try to say more on this topic in another place.

V

Non-Distributive Blameworthiness (NDB). We *govern* actions by authoring 'directives' (decisions, intentions, commands, etc.) that the actions execute. Any individual agent's action is *autonomous* or *self-governed* just if it is governed by him. More generally, for any actions $A_1, ..., A_n$, the property of being autonomous or self-governed is possessed by $A_1, ..., A_n$ just if they are governed by whoever performs them: it is *distributively* possessed by $A_1, ..., A_n$ just if, for each of $A_1, ..., A_n$, it is governed by whoever performs it; it is *non-distributively* possessed by $A_1, ..., A_n$ just if they are governed by whoever performs them, and it is not the case that, for each of $A_1, ..., A_n$, it is governed by whoever performs it; it is *collectively* possessed by $A_1, ..., A_n$ just if there is at least one instance of being governed by whoever performs them that is possessed by all of them (more simply, just if there is at least one instance of their *governance* by whoever performs them). When actions execute a collective decision made by their agents, they are collectively autonomous.

It is not obvious that collective autonomy entails non-distributive autonomy: sometimes, perhaps, when there is at least one instance of some actions' governance by their agents, there is also, for each of those actions, an instance of *its* governance by *its* agent, such that, for each of the actions, it is governed by whoever performs it.

I now stipulate that $a_1, ..., a_n$'s driving executes a collective decision to all drive at 80, which was made by them during the discussion period. Their acts of driving at 80 are, then, *collectively* autonomous or self-governed. My claim is that this renders them collectively blameworthy, that is, it makes it the case that there is one instance of blameworthiness that is possessed by them. Why so? Because these acts of driving are not only sub-optimal and impermissible but *expressive of their agents' collective will.*

A different example (adapted from Hayek 1973, p. 41) may help. Typically, a countryside track is created by each of a number of agents taking what seems to him to be the best path, and, typically, these acts of path-taking are distributively and not collectively autonomous: each is governed by its agent, and there is no instance of their governance by their agents. But we can imagine a contrast case in which a track is created by exactly the same means, that is, by each of some agents taking what seems to him to be the best path, but where these acts execute a collective decision made at a village

meeting, a decision to each take what seems to one to be the best path, in order that a track be made. The ensuing acts of path-taking are collectively autonomous, for the collective decision they execute is an instance of their governance by their agents.

In the typical case, there is something that some agents do, namely perform a plurality of path-taking acts—acts which collectively create a track—and in that sense, a plurality of acts attributable to them. But there is a way in which the *plurality* of acts are not fully *theirs*. For whilst the agents are, collectively, responsible for the acts, and the track that is their upshot, they did not *assume* responsibility for them, for they did not collectively *opt* to perform them. They lack the answerability that agents bear for actions that are *wholly* theirs, as actions to which agents have committed themselves are wholly theirs. But in the contrast case, it seems to me that, by collectively deciding to perform path-taking acts, the villagers assume collective ownership of, and responsibility for, these acts, and are appropriately identified with, asked to justify, and (as the case may be) praised or blamed for them, *because* they are, by virtue of their collective autonomy, expressive of their agents' collective will. I claim that the same is true of $a_1, ..., a_n$'s collectively autonomous acts of fast driving.

Of course, sometimes we find *non*-self-governed acts blameworthy. But—I hazard—these are seldom if ever judgements of the *collective* blameworthiness of acts performed by different agents. I suspect that this is so for two reasons. First, we tend to think of an individual agent as, like the long-distance driver of yore, *poised to snap* into self-governed action, should this become appropriate, as when he acts, or is on the verge of acting, impermissibly; hence, even when an agent is not governing himself, we will sometimes judge that he is blameworthy if he is, apparently, not sufficiently sensitive to the presence of a demand to snap into self-governed action. By contrast, pluralities of agents, even intercommunicative pluralities like $a_1, ..., a_n$, are not, as a matter of course, poised to snap into collectively self-governed action. Typically—if no one has authority or power over the others—agents need to talk over their options, reach a consensus, and make a collective decision if they are to collectively govern themselves: self-government is possible for pluralities, but is not *native* to them, as it is native to the individual, it must be negotiated and constructed by agents (each using what is native to him). Secondly, we sometimes judge that an individual

agent's impermissible but non-self-governed behaviour is blamewor-
thy because it is expressive of its agent's character (for example, his
mean or callous character, or simply his ill-disciplined, careless, un-
thinking character, his characteristic inability to govern himself
when appropriate). But we are seldom likewise inclined to think of
agents' impermissible activity as expressive of a character that they
collectively have.

I have argued that $a_1, \ldots, a_n$'s acts of driving at 80 are collectively
blameworthy (in part because they are collectively autonomous).
But it does not follow that they are non-distributively blameworthy.
Can it be shown that each act is not blameworthy?

This may appear doubtful, despite the fact that, as has been ar-
gued, each act is optimal and permissible. For what we now know is
that each of $a_1, \ldots, a_n$ has partaken of a collective decision that
renders their driving blameworthy. It may appear that each driver's
participation in this collective decision is correspondingly blame-
worthy, and that his driving inherits blameworthiness from his par-
ticipation in that decision.

Of course, there need be no question of any driver *governing* an
impermissible plurality of acts of driving at 80, as the drivers collec-
tively do. Each governs, at most, his own, permissible act, foreseeing
the rest, and gives the optimal response to what he foresees. But this
is not the whole story. Each driver is *party* to the drivers' governing
of the impermissible plurality of acts, and, in each case, this might
strike one as blameworthy, and as transmitting blameworthiness to
the agent's fast driving.

A final stipulation will remove the threat of distributive blame-
worthiness.[22] It is this: as every driver believes, and as their evidence
indicates, were they to use the discussion period to do *anything oth-
er* than make their collective decision to all drive at 80 (for example,
were they to make any other decision, or make no decision), each
would then choose his speed by coin-flipping, for each would be
neutral between his options, his perfect rationality and overriding
concern with road safety notwithstanding, as each would have no
belief and no evidence as to how others will drive. This augments
my earlier stipulation that this would all come to pass were $a_1, \ldots, a_n$

---

[22] If any reader worries that the introduction of the collective decision to drive at 80 should
prompt us to revise our earlier judgements that each driver drives optimally and permissi-
bly, is perfectly rational, and has an overriding concern with road safety, this final stipula-
tion should lay these concerns to rest.

to use the discussion period to make an agreement or collective deci-
sion to all drive at 60. Just to be clear, I am not amending my stipu-
lation that each of $a_1, ..., a_n$ believes that the others will drive at 80.
I am simply adding the further detail that, had they done anything
other than collectively decide to all drive at 80, this would not be the
case.

The collective decision is, then, something of a 'forced' delibera-
tive move, in that, as every driver believes, and as their evidence in-
dicates, for each driver, participation in the decision—be it
unconditional or conditional on others participating likewise—is
his only feasible means of ensuring that the drivers coordinate on
something other than the worst case, where drivers drive at varie-
gated speeds. For it is his only feasible means of deriving belief in,
and evidence for, a fact about everyone else's driving at a uniform
speed, belief which will determine, and evidence which will support,
his subsequent driving at the same speed. As no driver has any feasi-
ble means of ensuring that they coordinate on the best case, where
all drive at 60, and as each has overwhelming reason to avert the
worst case, I infer that, for each driver, participation in the decision
is—at the very least—non-blameworthy.

Should the 'forced' nature of the decision—the fact that it is the
only decision to drive at a uniform speed to which the drivers are
likely to stick—cause us to revise our earlier judgement that it
renders the drivers' driving collectively autonomous and thereby
collectively blameworthy?[23] I think not. For what 'forces' the deci-
sion is nothing but the dispositions of the drivers. Compare the indi-
vidual case: an individual's decisions can render him both self-
governing and answerable to others, even if, in the light of his dis-
positions, these decisions are 'forced' deliberative moves. If I decide
to miss a submission deadline, and this decision is 'forced' by my
disposition to not stick to any decision to meet my submission dead-
line, I govern myself all the same: my ownership of and responsibili-
ty for the action decided upon remain in place. I see no reason to
think that it is otherwise with collective decisions.

Why then is participation in the decision, and participation in its
execution *non*-blameworthy? Because, for each driver, dispositions
other than his own force him to participate. We saw that, when we
contemplate the permissibility of some action(s), performed by

---

[23] Thanks to David Wiggins for pressing a similar objection.

some agent(s), facts about the psychologies of some *other* agent(s) are reasonably conceived of as part of a 'fixed background', but facts about their *own* dispositions are not reasonably thus conceived. The same is true *mutatis mutandis* of blameworthiness. When we contemplate the blameworthiness of some action(s), facts about the dispositions of agents *other* than the agent(s) of said action(s) are reasonably conceived of as part of a 'fixed background', but facts about *their own* dispositions are not reasonably thus conceived. Hence, each driver's participation in their decision to all drive at 80 is excused by the others' dispositions to behave unpredictably were it not made, but the decision itself is not excused by the drivers' dispositions. It, and its execution, instantiate blameworthy self-government.[24]

VI

Can We Be Non-Distributively Blamed? I have argued that there can be actions $A_1, ..., A_n$ (with different individual agents) such that:

> (1) It is appropriate to blame whoever performed $A_1, ..., A_n$ for doing so.

> (2) It is not appropriate to blame whoever performed A_1 for doing so.

> ...

> (n+1) It is not appropriate to blame whoever performed A_n for doing so.

I think that a source of resistance to this thesis is the thought that the property of being blamed (i.e. of *receiving* blame, of having blame *laid* on you) can never be non-distributively had, i.e. the thought that there can be no agents $x_1, ..., x_n$ such that:

---

[24] One might think that each driver's driving inherits blameworthiness from his disposition to take himself to have no more reason to drive at one speed rather than the other, were a collective decision to drive at 80 not made. I reply as before. Even if such neutrality would be blameworthy, it does not obviously follow that each driver's *disposition* to be thus neutral is blameworthy, and it is, I think, highly doubtful that each driver's actual and otherwise non-blameworthy action of driving at 80 inherits any blameworthiness from a disposition of his that does not ever make itself manifest.

(1) $x_1, \ldots, x_n$ are blamed.

(2) x_1 is not blamed.

...

$(n+1)$ x_n is not blamed.

If that's right, then any possible case of the non-distributive blame-*worthiness* of actions with different individual agents is, thereby, a case of agents who cannot be blamed according to their worthiness; necessarily, either at least one non-blameworthy agent among them is blamed, or the blameworthy agents go unblamed.

This is not inconsistent with my thesis. But perhaps we should be wary of positing a form of blameworthiness that would systematically exclude the possibility that its parties reap, in blame, all and only what they sow in blameworthiness, baffling even God's best attempts to blame them in accordance with their worthiness of blame.

Any proof that agents can be non-distributively blamed would require a more thorough inquiry into the nature of blame than I can undertake here. In this section, I will simply try to show that two pregnant analogies between blame and other phenomena do not support the view that agents cannot be non-distributively blamed.

The first of the analogies is with punishment. Punitive sanctions are essentially harmful: I cannot penalize or punish if I do not thereby harm. Of course, some costs are easily absorbed, and some deprivations are scarcely noticed, but these can still be *harms*, and an attempt to punish that does not harm does not succeed. If, as seems plausible, there can be no non-distributively harmed agents (no harmed agents who are each unharmed) then, it may seem, there can be no non-distributively punished ones,[25] and so, if blame is just like punishment, no non-distributively blamed ones.

The first thing to say in response is that blame is not just like punishment. It is often harmful, but it is not essentially so.[26] I can blame someone—and express my blame to them—and not harm them, for example, if they simply do not care what I think or do.

But even if it were the case that we must harm whatever we blame, and so must harm a plurality if we blame them, and granting

---

[25] Jackson (1988, p. 264): '... you can't punish a group without punishing a goodly number of the members of the group'.

[26] Here I am influenced by Scanlon (1998, ch. 6; 2008, ch. 4) and Hieronymi (2004).

that we must, when we harm a plurality, harm at least one of them, it still does not follow that we must, when we blame a plurality, *blame* any one of them. After all, harm is not *sufficient* for blame.

We can grant that, as blame is *generally* harmful to its objects, and to harm a plurality one must harm at least one of them, NDB commonly results either in a non-blameworthy party receiving harm (because the parties are blamed), or in blameworthy parties not receiving the blame of which they are worthy. This may be a sorry consequence of NDB, but is scarcely an absurdity. After all, as harm is generally collaterally damaging, blameworthiness of *any* sort commonly results either in a non-blameworthy party receiving harm (because he is kith or kin to the blamed parties), or in blameworthy parties not receiving the blame of which they are worthy.

The second, stranger, analogy is with butter: (a) blame, like butter, comes in *amounts*: we say that x received much of the blame and y less of it, not that x received many of the blames, and y fewer of them; (b) blame is homeomerous: if x is *part* of some butter, then x is some butter;[27] likewise, if ever you get part of 'the blame', then what you get is blame; (c) we 'spread' blame much as we spread butter: to spread butter over, say, parsnips is, typically, to bestow, on the buttered vegetables, *portions* of some amount of butter; likewise, to 'spread' blame amongst agents is, typically, to bestow, on the blamed agents, *portions* of some (perhaps notional) amount of blame.

The analogy may encourage the thought that blame can only be spread amongst agents as butter is spread over parsnips, which is to say *distributed* amongst them, i.e. that necessarily, when agents are blamed—when blame is placed on them—each of them—or, if not each, one at least—is blamed, i.e. receives a portion of the blame, just as, when parsnips are buttered—when butter is placed on them—each of them—or if not each, one at least—is buttered, i.e. receives a portion of the butter.

But this exaggerates the analogy between placing blame on agents and placing butter on parsnips. For of course blame is not placed or put on things as butter is. To place blame on some agent(s) is either to *target* something—some act or attitude—at them, or to *give* something—some 'black mark' or 'booby prize', some *award* for

---

[27] Notice that I do not say 'If x is *a* part of some butter, then x is butter', for butter contains electrons.

outstandingly poor achievement—to them. Now, an act or attitude can be targeted at a plurality without it, any part or portion of it, or anything of the same kind as it, being targeted at any one of them: an actor who fears his audience targets his fear at them, but may not target it, or any part or portion of it, or any act or attitude of the same kind as it, at any one of his audience. Likewise, an award can be given to a plurality without it, or any part or portion of it, or anything of the same kind as it, being given to any one of them. Sometimes, perhaps, a single award goes to *each* of a plurality, as when a tie occurs. But when the achievement recognized has been done in collaboration (one thinks of Oscars and Nobel prizes), an award will go to a plurality, none of whom receive it, any part or portion of it, or anything of the same kind as it.[28]

In short, blame can be spread, and is like *a* spread, but we receive it, not as parsnips receive butter, but as agents come by targeted acts or attitudes and/or awards.

VII

Corporate Blameworthiness. Call an agent that is exhaustively constituted by, or composed of, a plurality of agents a *corporate agent*. This paper presupposes no answer to these questions: (i) Are there, or could there be, corporate agents? (ii) Are there, or could there be, blameworthy corporate agents? My thesis is that there can be severally non-blameworthy agents who are themselves blameworthy, and not that there can be severally non-blameworthy agents who constitute or compose something else that is blameworthy.

A corollary of my thesis is that a tempting form of argument for corporate blameworthiness is weak. Informally stated, it is this: blame can be called for by the actions of agents who each act blamelessly; when this occurs, there must exist some *other* agent—acting, as it were, behind the scenes—who is able to 'take the blame'; and the best hypothesis is that this agent is an incorporation of the blameless agents.

---

[28] Those who non-distributively receive an award typically each receive a trophy; the trophies perform the useful function of *marking* the presentation of the award. Likewise, perhaps, when agents are non-distributively blamed, their being thus blamed is, typically, recorded and represented (*ceremonialized*, even), by our acting towards *each* of them in certain sorts of (generally but inessentially harmful) ways.

More precisely, it is this: there can be actions $A_1, ..., A_n$, with different human individuals as their agents, such that:

(1)   It is appropriate to blame *something* for performing A_1, ..., A_n.

(2)   It is not appropriate to blame A_1's human agent for performing it.

...

($n+1$) It is not appropriate to blame A_n's human agent for performing it.

Therefore, by inference to the best explanation of (1), given (2), ..., ($n+1$):

(C)   There is an agent x that is exhaustively constituted by, or composed of, the human agents of $A_1, ..., A_n$, such that x is the 'distal' agent of $A_1, ..., A_n$, and such that it is appropriate to blame x for performing $A_1, ..., A_n$.

This is a poor argument, for, given the possibility of NDB, and absent independent evidence for the existence of an agent x that incorporates the agents of $A_1, ..., A_n$ and performs their actions, there is a better explanation of (1) than (C), which is consistent with (2), ..., ($n+1$), namely that it is appropriate to blame the human agents of $A_1, ..., A_n$ for performing them. This is (a) a less *ad hoc* explanation than (C), as it identifies the blameworthy 'something' with the 'proximal' human agents of the actions that render blame appropriate, rather than with something *else*,[29] and (b) more parsimonious than (C), as it posits the existence of no agent other than those agents.[30] So, (C) is not the *best* explanation, and abductive inference to it is invalid.[31]

---

[29] Let no one doubt that something can be *some things*. If something I ate last night has given me food poisoning, this does not exclude the possibility that some prawns have given me food poisoning, even though no prawn has. Of course, 'something' is *grammatically* singular. But grammatical and semantical number sometimes diverge, just as grammatical and semantical gender do: some occurrences of 'pants' are no more *semantically* plural than is 'HMS *Victory*' *semantically* female.

[30] It may also be a morally superior hypothesis. (C) seems to, as it is said, 'let the individuals off the hook'; on the rival hypothesis, the individuals are on the hook (even though no one of them is).

[31] Thanks to the London audience, also to audiences in Manchester, Lancaster and Cologne, and to Oliver Black, Richard Evans, Harry Lesser, David Liggins, Fraser MacBride, Mike Martin and Lee Walters.

Philosophy
Arthur Lewis Building
University of Manchester
Oxford Road
Manchester M13 9PL
UK
thomas.smith@manchester.ac.uk

REFERENCES

Bacharach, Michael 2006: *Beyond Individual Choice: Teams and Frames in Game Theory*, ed. Natalie Gold and Robert Sugden. Princeton, NJ: Princeton University Press.

Black, Oliver 2004: 'Agreements, Undertakings, and Practical Reason'. *Legal Theory*, 10, pp. 77–95.

——2007: 'Two Theories of Agreement'. *Legal Theory*, 13, pp. 1–22.

Dancy, Jonathan 2000: *Practical Reality*. Oxford: Oxford University Press.

Feldman, Fred 1986: *Doing the Best We Can: An Essay in Informal Deontic Logic*. Dordrecht, Netherlands: D. Reidel.

Gilbert, Margaret 1996: *Living Together: Rationality, Sociality and Obligation*. Lanham, MD: Rowman and Littlefield.

Hayek, F. A. 1973: Selection from 'Scientism and the Study of Society'. In John O'Neill (ed.), *Modes of Individualism and Collectivism*. London: Heinemann.

Hieronymi, Pamela 2004: 'The Force and Fairness of Blame'. *Philosophical Perspectives*, 18, pp. 115–48.

Humberstone, I. L. 1983: 'The Background of Circumstances'. *Pacific Philosophical Quarterly*, 64, pp. 19–34.

——1991: 'Two Kinds of Agent-Relativity'. *Philosophical Quarterly*, 41, pp. 144–66.

Jackson, Frank 1987: 'Group Morality'. In Philip Pettit, Richard Sylvan and Jean Norman (eds.), *Metaphysics and Morality: Essays in Honour of J. J. C. Smart*, pp. 92–110. Oxford: Blackwell.

——1988: 'Understanding the Logic of Obligation'. *Proceedings of the Aristotelian Society Supplementary Volume 62*, pp. 255–70.

Keynes, John Neville 1906: *Studies and Exercises in Formal Logic*. London: Macmillan.

Lewis, David 1973: *Counterfactuals*. Oxford: Blackwell.

Oliver, Alex, and Timothy Smiley 2001: 'Strategies for a Logic of Plurals'. *Philosophical Quarterly*, 51, pp. 289–306.

Parfit, Derek 1988: 'What We Together Do'. Unpublished MS available at

<http://individual.utoronto.ca/stafforini/parfit/parfit_-_what_we_together_do.pdf>

——1989: *Reasons and Persons*. Oxford: Clarendon.

——2003: 'Justifiability to Each Person'. *Ratio*, 16(4), pp. 368–90.

——forthcoming: *Climbing the Mountain*. Draft at <http://individual.utoronto.ca/stafforini/parfit/parfit_-_climbing_the_mountain. pdf>.

Pink, Thomas 2004: 'Moral Obligation'. *Royal Institute of Philosophy Supplement*, 54, pp. 159–85.

Raz, Joseph 1999: *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press.

Regan, Donald 1980: *Utilitarianism and Cooperation*. Oxford: Clarendon.

Roth, Abraham Sesshu 2004: 'Shared Agency and Contralateral Commitments'. *Philosophical Review*, 113, pp. 359–410.

Scanlon, T. M. 1998: *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

——2003: 'Replies'. *Ratio*, 16(4), pp. 424–39.

——2008: *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Harvard University Press.

Schelling, Thomas C. 1960: *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Smith, Thomas H. 2009a: 'Intentional Joint Activity'. Unpublished MS.

——2009b: 'Meetings of Minds'. Unpublished MS.

Velleman, J. David 1997: 'How to Share an Intention'. *Philosophy and Phenomenological Research*, 57, pp. 29–50.