

Forthcoming in *Oxford Readings in Epistemology (Special Theme: Social Epistemology)*, vol. 3, eds. Tamar Gendler and John Hawthorne (2009)

Can There Be Institutional Virtues?

In natural speech we often attribute virtues and vices not only to individuals but also to groups of various kinds. We say things like ‘The jury is fair-minded’, ‘The research team is tenacious’, ‘The appointments committee is scrupulous’, or indeed ‘The Metropolitan Police is institutionally racist’. But what exactly are we doing when we say these things? This question incorporates the more general metaphysical question about the nature of groups—whether a group is to be understood as nothing more than a sum of individuals, or whether we also need in our repertoire a concept of a group as a collective which, though composed of individuals, cannot be reduced to the sum of its parts. My argument will be in line with those who argue the latter, and I shall make core use of Margaret Gilbert’s conception of collectives as ‘plural subjects’ in my construction of collective virtue.¹ That groups can display virtues as collectives is of course commensurate with the fact that they can also display virtues simply as clusters of virtuous individuals. The range of possibilities in play reflects partly the effective ambiguity of ‘group’, and partly the fact that ‘virtue’ too may be understood in at least two ways: in a motivationally demanding way, as in the Aristotelian conception; or as a sheer excellence or skill, as in the Platonic and Stoic conception.² I

¹ For Margaret Gilbert’s seminal account of social groups as ‘plural subjects’ see *On Social Facts* (Princeton, New Jersey: Princeton University Press, 1989). See also *Sociality and Responsibility: New Essays in Plural Subject Theory* (Rowman and Littlefield, Lanham, MD, 2000).

² For an instructive discussion of these contrasting conceptions, see Julia Annas, ‘The Structure of Virtue’, Zagzebski and DePaul eds. *Intellectual Virtue* (Oxford: OUP, 2003).

will argue for a conception of institutional virtue that works for both kinds of virtue, and both senses of ‘group’—the relevant philosophical challenge of course being in respect of the collective sense of group. Once the model for collective virtue is established from the argument of the first two sections, I go on in the final section to explain the respects in which institutions may or may not be literally described as displaying virtues, or indeed vices.

1. Institutionalized Groups—Three Aspects

There are many senses of social group, ranging from the merely statistical (the group of people who drive Swedish cars) to the explicitly self-identified (as when someone speaks as a member of a given social group—‘Speaking as a gay man, I...’). The sense of group that is relevant for my purposes is any that sustains group agency, for my interest is in the possibility of appraising group conduct, including epistemic conduct, in terms of virtue and vice. Sometimes such groups are entirely informally constituted, as in Gilbert’s familiar example ‘Let’s go for a walk’, and off they go. But often their constitution and conduct are at least partly institutionalized, so that they are at least loosely governed by a set of institutional norms or procedures. Teams, committees, juries, governments are all examples of more or less institutionalized groups, and the dimension of institutionalization means that such groups may be described under three different aspects:

- 1) a number of individuals (the group considered as the sum of its component individuals)

- 2) a collective (the group considered as non-reducible to its component individuals)
- 3) an institutional structure (its formal and procedural structure)

If, for instance, a member of the jury in a contentious court case were to privately remark, ‘The jury was fair-minded’, she might quite properly be describing the jury under either of the first two aspects. That is, she might be saying (1) that enough of the *individual* jurors displayed fair-mindedness; or she might be saying (2) that the jury taken *collectively* displayed fair-mindedness. (Or, in case it is indeterminate which of these things is being said, let us simply say that either fact would render her statement true.) If, following up with a more general reflection, she were to add ‘The jury is a just system’, she would be referring to the institution of trial-by-jury defined purely structurally and procedurally in the manner specified in (3)—that is, in abstraction from the performance of any particular collective jury or number of jurors who realize the institutional procedures in any given instance.

I will discuss institutions *per se* in the third section of this paper. For the moment, however, let us ask what we are doing when we say of a jury that it was fair-minded. This concerns only the first and second aspects. I take it for granted that there can be instances of groups displaying features in the manner of (1)—‘People who choose Swedish cars are safety conscious’. The immediate task is rather to vindicate the idea that groups can display features and attributes in the manner specified in (2)—that is, as collectives. The fundamental controversy in the metaphysics of groups is the argument between summative and non-summative accounts of groups.

Summative accounts analyze groups exclusively as per the first aspect given above, as

a number of individuals. Non-summative accounts insist that there need be no such reduction, and that groups, though composed of individuals, may be more than (or, at any rate, other than) the sum of their parts. According to a basic summative account, if someone says ‘The jury is fair-minded’, this must be understood as the claim that enough of the jurors are fair-minded individuals. Now, as others have argued in respect of other sorts of group feature, this cannot be right. Summativism does not work as a general account of group features, for there can be cases where a group possesses a feature that few or even none of its component individuals possess (so individual possession of the feature is not necessary); and there can be cases where the group lacks a feature even though it is possessed by many or even all of the component individuals (so individual possession of the feature is not sufficient).³

One style of example offered against summativism about group features plays on a thoroughly contingent, non-normative type of influence that the social context of the group can have on what features people display. Social context can bring all sorts of pressures to bear on how we behave, how we think, and what judgements we make—sometimes these pressures may be to good effect, sometimes bad. For instance, it might be pointed out that there could be a football team that was competitive even while none of its members were competitive individuals, indeed they are all distinctly uncompetitive souls. But when they get together as players for their local club, luckily they come to care about winning in a manner that generates an impassioned, if highly localized, collective competitiveness at team-level. This sort of

³ For Gilbert’s original case against summativism in respect of group belief, see Gilbert (1989), ch.5, esp. 257-60. For more recent statements, see for instance ‘Collective Epistemology’, *Episteme* Vol. 1, No. 2 (October 2004), 95-107; and ‘Belief and Acceptance as Features of Groups’, *Protosociology* Vol. 16 (2002), 35-69.

example has some impetus in the direction of non-summativism, since it at least shows that for a group to have a given feature (competitiveness) it is not necessary for the component individuals to have that feature in general; and it also shows that for a group to have a given feature (uncompetitiveness) it is not sufficient that its members should have it as individuals. Reza Lahroodi has put forward a suggestive example that I would categorize as being of this type, though where the effect of the contextual pressures happens to be bad. He imagines an administrative church committee made up of individual members who are each open-minded about gay rights. But unfortunately when they get together in the context of the church committee a certain closed-mindedness comes over them, perhaps because they ‘want others to think they are towing the church line on this issue. They may clam up in the presence of other members if they anticipate negative reaction by powerful authorities outside the group’.⁴

While examples such as this negate ‘simple summativism’, defined as the view that ‘a group G has the trait T if, and only if, all or most members of G have T’⁵, nonetheless they cannot clinch the argument against the individualistic spirit of summativism, for one kind of less simple summativist can surely reply that even if the individuals (to return to the football example) lack competitiveness in other contexts, still the team’s competitiveness is to be understood as none other than the sum of the

⁴ Reza Lahroodi, ‘Collective Epistemic Virtues’, *Social Epistemology*, 21:3, 281-297; 288. This is only one of two different sorts of example he presents against summativism. The second concerns the possibility of two committees with identical membership nonetheless having non-identical features owing to the differing remits of the committees. I would categorize this as falling under what I label below as the ‘practical identities’ style of example.

⁵ Lahroodi, *op.cit.*; 285.

individuals' contributions of competitiveness—the fact that it is only in the group context that each displays the attribute is merely to note that the feature in question is out of character for each member. The summativist can therefore maintain that the group feature is simply the sum of the individual features, and a summativist reduction beckons.

A more decisive style of counter-example to summativism will not, I suggest, turn on quite such a contingent type of influence as this, but rather on the normative and psychologically structured influence that a subject's various 'practical identities'⁶ have on what features and attributes she is committed to displaying. The mere fact that social subjects have a range of practical identities (so that one may confront a situation, decision, or choice *as* a professional, *as* a parent, *as* a friend, *as* a gay man, *as* a Christian, *as* an interested or disinterested party, and so on) means that there can be tension, and sometimes downright conflict, between the commitments associated with different practical identities of the same person. This in turn generates the possibility of that individual having a certain attribute only *as* a group-member and not as a private individual. For instance, in order for the committee of the local amateur dramatics society to care deeply about the society's prospects it is not necessary that the individual committee members, *qua* private individuals, care deeply. It may be that as a local resident who is regularly inconvenienced by the society's occupying the town hall and taking up all the parking places on a Thursday evening, one would not shed a tear if the society ceased to operate. None the less, *qua* member of the society one is committed to engaging whole-heartedly in the

⁶ Christine Korsgaard coins this useful term in *The Sources of Normativity* (Cambridge University Press, 1996); see chapter 3.

endeavour to see it survive and flourish. Indeed all the society's members may have their own version of similarly conflicting feelings and commitments, but it is nonetheless clear to each member which among their various attitudes and motivations they are obliged to bring to bear in any given context, depending on which hat they are wearing.

This style of example, exploited by Gilbert most notably in her account of group belief, tells more decisively against the summativist approach to group features. It illustrates how a group can have a feature or attribute (whole-hearted commitment to the society's future) which is lacked by every member *qua* private individual. And it also illustrates how every member of a group's having a given feature (not caring deeply about the society's future) is insufficient for the group to have it. Here the summativist cannot respond by saying that the group feature is to be understood as the sum of the individual features, for the individual features in question (each member's wholehearted commitment *qua* committee member) are not found at the level of individuals considered independently from the group, for the individuals only have that feature if they are wearing their group-member hat. Some practical identities of individuals are thus intrinsically group-involving, and in such cases there is no lower level of group-*independent* features to which the higher-level features can be reduced. Any attempted reduction of the group to a sum of uncommitted non-group-identified individuals would literally change the subject, and so fail.

Finally, there is a third species of example to mention, because once again it blocks any summativist reduction, but in a different way from the previous case. Once again, the group feature is not found at the individual level, but here the higher-level

feature is generated by a kind of invisible hand. Imagine a debating society, the members of which are thoroughly prejudiced individuals. But their prejudices are all opposed and equally balanced, so that they cancel each other out in debate and the debate overall displays not prejudice but rather neutrality. (Compare the adversarial structure and style of argument that is proper to the battle between prosecution and defence in a courtroom.) Such a debating society, then, is itself non-prejudiced, even while all of its members, and their contributions to the group's activities, are prejudiced. The higher-level feature is nowhere to be found at the lower level, and is explained by the way in which the individual features synthesize to create a quite different feature at group level. Unlike the 'practical identities' style of example, examples of the 'invisible hand' variety do permit a kind of reduction: the lower level is explanatorily basic. But the reduction is not of the summativist kind, for there can be no simple aggregation of lower-level features that adds up to the higher-level feature. Thus the 'invisible hand' case and the 'practical identities' case represent two principal ways in which non-summative group-level features are generated.

So far I have made reference to group features in general. A virtue is a kind of feature, but it may be that virtue is a special sort of feature when it comes to groups. Let's look at the example briefly mentioned earlier. Imagine a jury whose conduct in a prolonged court case renders it deserving of the description 'fair-minded', even while few of the individual jurors merit that description either in general as individuals, or in his or her particular contribution to the deliberations of this jury. So far so good; we can imagine that the jury as a whole might yet exhibit fair-mindedness. But I said *few* of the jurors; what if *none* of the jurors made any kind of fair-minded individual contribution? Perhaps, like the members of our debating

society, these jurors are a frankly prejudiced lot, but diversely so, with the happy result that their prejudices all cancel each other out as they move collectively towards their verdict. Thus the invisible hand. But here we see that virtue is a rather special sort of group attribute. So long as we consider it a requirement of virtue that the good conduct is performed *because of* the good motive or skill—as opposed to a misunderstanding or lucky accident—so that we can credit the agent (in this case a group agent) for that conduct, then it seems too much of a stretch to suppose there could be a fair-minded jury none of whose members displayed any fair-mindedness. For the fair-mindedness exhibited here is not achieved because of any group sensitivity to the demands of fairness—it is a mere accident of the way the individuals’ prejudices happen to cancel one another out. Accordingly, there is no sense in which the jury is creditworthy for its conduct. In such a case as this, the group’s behaviour might display something like group fair-mindedness, but it would be too much of a stretch to say it possessed the virtue. The same point applies to vice, which mirrors virtue in requiring that the bad action be done because of the bad motive or deficit of skill. Like virtuous action, vicious action must flow from the agent’s character. Invisible hand examples generate group features as mere by-products, too epiphenomenal to qualify as genuine group virtue or vice; rather they exemplify how groups may sometimes come to possess shadow versions of virtues and vices, so that if we loosely apply those concepts, we do so without the normal implications of praise or blame.

Nothing else hangs on where one is inclined to draw the line on this, however. For whether or not one agrees that the special nature of virtue and vice renders them inapplicable to invisible hand examples, still the practical identities style of example

does the job, and so the case against summativism as a general account of group features and attributes, including virtues and vices, is independently established. Sometimes group features may be a matter of a sufficient number of members displaying the trait as individuals (as in our example of buyers of Swedish cars being safety conscious); but sometimes they are irreducibly collective. So how should we model group virtue and vice of the collective kind? A non-summative account is called for, and I hope to show that the model given by Gilbert in her classic notion of a ‘plural subject’—originally applied principally to group action and group belief—provides an excellent template for our thinking about group virtue and vice of the collective sort.

2. Collective Virtue as Plural Virtue (Introducing Passengers and Stowaways)

Gilbert’s notion of a plural subject is such that one is constituted whenever some number of individuals jointly commit, under conditions of common knowledge, to a given action, belief, or goal:

In order for two people to share in an action each must have made clear to the other that he is willing to join forces with him in accepting the relevant goal... each must manifest his willingness to constitute with the other a plural subject of the goal in question.⁷

⁷ Gilbert, *On Social Facts*; 204.

As she has formulated it in later work, the individuals commit to ‘X-ing as a body’, and the joint commitment might be practical or cognitive. Looking ahead momentarily to the idea that plural subjects might possess virtues, we can envisage how, in the case of intellectual virtues, the joint commitment would combine both practical and cognitive elements.

Gilbert’s notion of ‘commitment’ here is such that if someone reneges on a commitment to X-ing as a body, he is subject to some kind of rebuke, or at least puzzlement. One owes one’s fellow walkers an explanation if, having committed to go for a walk together, one suddenly jumps on a passing bus. I think this captures the psychology and the ethics of joint commitment, but it is especially apt in the case of collective virtue. Imagine the members of a science research team have jointly committed to diligence and thoroughness in respect of a certain range of findings (they commit to re-trying the experiments one more time, just to be absolutely sure); but on Monday morning one of them is seen to have moved on to work on something else instead. His co-researchers would *hold* him to his seemingly forgotten commitment by demanding some sort of explanation, at the very least. The appropriateness of rebuke, or at least a demand for explanation, is what gives content to the notion of commitment. A commitment is something others are entitled to hold us to, and this registers its ethical dimension; but where it is specifically a commitment of virtue on which we have reneged, the rebuke has a double ethical aspect.

I propose that Gilbert’s notion of a plural subject provides a template for group virtue—‘plural virtue’ as we might call it—whether we have in mind a motive-

based model of virtue or a skill-based model. I hope to show that, if appropriately adapted, it can serve as a way of pluralizing either model of virtue, and so my argument for the possibility of collective virtue can remain neutral with respect to the controversy over how to conceive of virtue in general. Actually, I think there is a strong case for taking the motive-based and skill-based conceptions of virtue not as competing and exclusive, but rather as applying to different kinds of virtue. Some virtues, after all, seem to lend themselves more naturally to a motive-based interpretation (kindness, compassion, charity, generosity) because what is paramount in these virtues is the agent's attitude; while others, by contrast, seem to lend themselves more naturally to a skill-based interpretation (vigilance, honesty, justice, inventiveness) because what is paramount is the agent's performance. Inasmuch as this is so, there is a further motivation to come up with a model of collective virtue that excludes neither conception.

First, let us consider how to construct collective motive-based virtue. What is needed is to employ the notion of joint commitment in respect of group *motive*. If, under conditions of common knowledge, a number of individuals commit to a virtuous moral or epistemic motive, they thereby constitute themselves as the plural subject of that *collective motive*. Joint commitment to a motive should be understood here as shorthand for a joint commitment to achieving the good end of the motive *because* it is good. Or, as we might put it, a joint commitment to a virtuous motive is a matter of jointly committing to the virtuous end for the right reason. Note that group members need not possess the motive *as* individuals. Rather, in jointly committing to it, they each come to possess it *qua* member of the group. Collective motive is in this respect analogous to collective belief—both are modelled according to our ‘practical

identities' style of example in that there is room for considerable tension between one's personal motives/beliefs and the motives/beliefs one comes to have in virtue of this or that group identity. Let us now add to this group motive the requisite reliability condition⁸; and *voilà*, we have a collective virtue.

Consider our diligent and thorough research team. Its members all jointly commit to the motives of diligence and thoroughness; and the team lives up to those motives by proving reliable, over an appropriate span of time and contexts, in achieving their ends. This research team displays irreducibly collective forms of diligence and thoroughness. They display these virtues 'as a body', or *as one*, where this is correlated with the fact that the team's answerability to these high standards of scientific inquiry is similarly collective—praise and blame for the team's conduct is directed at the team considered as a collective. (Only secondarily might we—perhaps on discovering that certain members made a greater contribution than others—come also to attribute individual praise to reflect the different levels of contribution.) Exactly the same account works for vice if we substitute a bad motive: imagine a research team whose members are jointly motivated to presenting their findings as proven no matter what (even if it takes falsifying the empirical evidence).

There are, not surprisingly in this somewhat off-the-peg account of collective virtue, a couple of ill-placed creases that need ironing out. But first let us deal with a conspicuous if superficial wrinkle: namely, the worry that any such account entails a

⁸ Here I can remain neutral once again in relation to a controversy in virtue theory—this time the controversy over whether the reliability must be in respect of achieving the immediate or the ultimate end of the virtuous motive. (See Annas, *op.cit.*)

commitment to group minds of a metaphysically spooky sort. Lahroodi expresses this worry in relation not to group motive but to the idea of a group *faculty*: ‘The commitment to group minds would be unavoidable should virtues, construed as cognitive *faculties*, be attributed to groups’.⁹ And for this reason he conducts his own inquiry into collective (epistemic) virtue according to a conception of virtue that makes no mention of faculties but only traits, the idea of group traits striking him as metaphysically unspooky.

But I think he needn’t have worried about the idea of group faculties. Gilbert’s model makes clear how group commitments are a special sort of combination of individual commitments. It is the *pooling* of wills, or other resources such as faculties, under conditions of common knowledge that does the pluralizing work. The notion of a collective faculty would, on this model, come to nothing more metaphysically spooky than pooled faculties, which is to say, nothing metaphysically spooky at all. When wills are pooled, they generate an *as-if* group will: ‘these wills will be directed at that end, *as if* they belonged to a single person. That is, the coherence of the behaviour which is their output will approximate in coherence to the output of the will of a single person acting in pursuit of a goal of his own’.¹⁰ And so we can say that when faculties are pooled, they generate an *as-if* group faculty. This as-if group faculty might take various forms, since the notion of pooling leaves it open exactly how the faculties might be combined. But one natural and efficient form of group faculty involves the division of labour. Imagine a night watch team, consisting of four soldiers. They might sensibly pool their faculties by dividing the labour so that one

⁹ Lahroodi, *op.cit.*; 284.

¹⁰ Gilbert, *On Social Facts*; 211 (*italics added*).

looks North, another South, another East and the other West. Alternatively, and presumably less efficiently, they might pool their faculties by each of them constantly shifting their gaze to look every which way. If the night watch team aims to display the virtue of vigilance, its members will jointly commit to the more reliable division of labour. Need we attribute any virtuous *motives* to these soldiers? I don't think so. Perhaps they just *do* it—that is how they were trained, and it never occurs to any of them to proceed in any other way than what is in fact the more reliable, indeed collectively skilful way. Here we come to a natural use for the excellence or skill-based model of virtue. One might rightly attribute the virtue of vigilance to this night watch team, where that is to be construed as a sheer collective skill, with no particular collective motivational implications. That's all you want in a night watch, after all—never mind the motivations, what matters is plain performance.

So what should we say about the kind of collective virtue displayed by this night watch? Here we have no particular use for joint commitment in respect of motive, but only in respect of ends. The team members jointly commit to achieving the end of vigilance by way of a good method (they commit to staying on the alert for enemy movement by way of an efficient division of labour). If we now add to this collective skill that the watch does indeed reliably achieve its end; then *voilà*, we have the collective excellence of vigilance. Suppose now, with vice in mind as a deficit of skill, we imagine the next night watch team as a hopeless bunch of slackers—they wile the night away telling jokes, playing cards, nodding off, and in one or another manner signally failing to jointly commit to the end of vigilance. Given that vigilance and negligence are exclusive opposites for a night watch, the watch thereby displays the collective vice of negligence. And so we see that in connection with a skill-based

conception of virtue and vice, Gilbert's idea of plural subjectivity seems readily able to handle the demands of collective virtue. It provides for a model of collective virtue which seems to work for both motive-based and skill-based conceptions of virtue and vice.

But at this point we must turn our attention to the two ill-placed creases in our new garment, and see if we can iron them out. I began the discussion of group virtue with the anti-summativist thought that a group can possess a virtue, or indeed a vice, in spite of the majority of its members lacking it as individuals. This is a crucially important point not only from the perspective of the philosophy of collectives but also from the first order ethical-political point of view of identifying those collective practices of racism and other kinds of discrimination or prejudice that are not a matter of individuals' personal attitudes. When the London Metropolitan police was diagnosed as 'institutionally racist' there was much confused public discussion around the question whether this boiled down to an accusation of individual police officers being racists.¹¹ This is evidenced in the publicly expressed advance fear on

¹¹ The accusation of institutional racism was made in the Macpherson report on the police handling of the murder of Stephen Lawrence, a black teenager fatally stabbed in London by a small gang of white teenagers on 22nd April 1993. The report, headed by Sir William Macpherson of Cluny, deals with racism principally in chapter 6, and places great emphasis on the distinction between individual racist attitudes and unintentional racism of outcomes in terms of how black and ethnic minorities are treated by the police. The report was published in 1999 and can be found at <http://www.archive.official-documents.co.uk/document/cm42/4262/4262.htm>. At 6.34 it defines for the purposes of the report the concept of institutional racism as follows: 'The collective failure of an organisation to provide an appropriate and professional service to people because of their colour, culture, or ethnic origin. It can be seen or detected in processes, attitudes and behaviour which amount to discrimination through unwitting prejudice, ignorance, thoughtlessness and racist stereotyping which disadvantage minority ethnic people.' The paragraph continues, 'It persists because of the failure of the organisation openly and adequately to recognise and address its existence and causes by policy, example and leadership. Without recognition and

the part of the then Commissioner of the Metropolitan Police Service that if the report were to label the MPS institutionally racist, then the average police officer and member of the public would understand the accusation to be that the majority of police officers ‘go about their daily lives with racism in their minds and in their endeavour’.¹² The possibility of continued public misunderstanding of the role that police racism had played in the case, and the great pains made in the Macpherson report to improve that understanding, underline the social importance of the philosophical distinction between attitudes held at the level of the individual and attitudes held at the level of the group, a group which may sustain, for instance, a certain racist ‘canteen culture’.¹³

Regardless of whether it was true that any number of individual officers were indeed personally racist, then, it was crucial to the content of the claim of institutional racism that it did not entail it. Accordingly, our philosophical picture of collective virtue and vice has been inspired in part by this anti-summative impetus. But now that we have presented collective virtue and vice as plural virtue and vice, using Gilbert’s model of plural subjects, we should worry that the conception as so far advocated might not seem quite so *collective* as we had imagined. For the trouble with modelling collective virtue as Gilbertian joint commitment is that this requires

action to eliminate such racism it can prevail as part of the ethos or culture of the organisation. It is a corrosive disease.’

Recognizing the phenomenon of course depends upon appreciating that institutional racism does not require racist attitudes on the part of individuals.

¹² He is quoted in the Macpherson report at 6.46.

¹³ On this point of ‘canteen culture’ or ‘occupational culture’, see paragraph 6.28 of the Macpherson report, which quotes oral evidence given by the three representatives of the MPS Black Police Association.

conditions of common knowledge; and the requirement of common knowledge seems to entail both that the commitment is *unanimous* in the group, and that members are *self-aware* in their commitment. On the face of it, neither of these features of joint commitment fits with what we should be saying about collective virtue: the unanimity requirement conflicts with our anti-summativist inspiration which would allow that groups can have virtues or vices even if the majority of their members lack them; and the self-awareness requirement conflicts with an essential feature of virtue and vice generally, namely that the subject of one need not be aware of possessing it. A compassionate person need not be aware that he is compassionate; an intellectually courageous person need not be aware that she is intellectually courageous; indeed it is famously plausible that a modest person is necessarily unaware that he is modest.¹⁴

Lahroodi has rightly emphasized the problem of the self-awareness requirement as a difficulty for a Gilbert-styled joint commitment model of collective virtue.¹⁵ But we should give equal emphasis to the distinct requirement of unanimity. Can we overcome these two problems? Let us examine each more closely. First, the self-awareness requirement. There are two points to be made in this connection:

(1) In the case of individual virtue, the fact that it does not require self-awareness is partly explained by the possibility that the subject may not *conceive* of her good motive or skill *as* virtuous. Indeed she may not consciously conceptualize it *as* anything much, but rather she may act spontaneously from habit, her focus being

¹⁴ For a discussion of such ‘virtues of ignorance’ see Julia Driver’s *Uneasy Virtue* (Cambridge: Cambridge University Press, 2001), chapter 2.

¹⁵ Lahroodi, *op.cit.*; 292.

simply on achieving the good end of that motive or skill. Now this can be true also of collectives. There is no reason to say that a vigilant night watch (to invoke our skill-based example) or a fair-minded jury (considered as a motive-based example) need be aware of its virtue, because the group might not conceive of the jointly committed good motive or skill *as* virtuous. When the individuals all jointly commit to the common end (of vigilance, or fair-mindedness) they might simply have that end in mind without conceptualizing it *as* anything. Either way, group members need not be aware of the virtuous nature of their jointly committed motives or skills.

(2) In the case of individual virtue, the subject may not be aware of his virtue because he might not be aware that he is *reliable* in achieving the relevant end of the good motive or skill. This too is so of collectives. Neither our vigilant night watch team nor our fair-minded jury need be aware that it is reliable in achieving the relevant end. They might just know they are doing their best, or doing their job, or something of that sort. For this reason too, a collective need not be aware that it possesses a virtue any more than an individual need be so aware; and for precisely the same reasons.

The general point to take from (1) and (2) is that even while all parties must indeed be aware of the joint commitment to a given motive or skill, it does not follow from this that they are aware of all features of that to which they are committed (such as that it is virtuous), or again of their performance in relation to it (such as that it achieves reliability). Either of these is independently sufficient for the group to be unaware that it possesses a virtue or vice. This shows that what appeared to be a serious problem for anyone advancing a Gilbert-styled joint commitment model of

collective virtue and vice, is not after all a problem. Indeed collective virtue on this model simply mirrors individual virtue in relation to self-awareness.

What about the second of our two problems—the unanimity requirement? I propose there are at least *two* ways a group may earn the title of virtue without all individual members possessing the good motive or skill. Both possibilities exploit that feature of joint commitment that figures large in Gilbert’s account of group belief: the fact that a person has more than one practical identity, so that with one hat on they believe that *p*, while with another hat on, they do not believe that *p*, or even believe that *not-p*.¹⁶ (On this model of group belief we might say that jointly committed individual acceptances that *p* can generate a group belief that *p*, though I am uncertain whether Gilbert would endorse this way of putting it.) There are two possibilities I want to propose for group virtue without unanimous possession of the good motive/skill.

Firstly, imagine a group—whether a plural subject or some other kind of group, such as a committee so far constituted only by institutional fiat (they may not have got down to any collective business yet)—and this group contains a sub-group that is a plural subject. Imagine that the members of this sub-group possess the virtuous motive or skill as individuals, and jointly commit to it, with the other members of the greater group jointly committing only as *passengers*—they lack the good motive/skill as individuals but ‘go along with’ it *qua* group members (minimally acting according to it). If we add that the group as a whole reliably achieves the end

¹⁶ Gilbert does not herself put it in terms of a shift of ‘practical identity’, but I believe that is what is at stake.

of the motive/skill, then the group as a whole possesses the relevant virtue.¹⁷ Note that this possibility exploits the fact that in Gilbert's account of group belief there are effectively two routes that an individual may take to a joint commitment to believing *p* as a body: (i) via her own personal belief that *p*; and (ii) via a mere 'going along with' or 'letting it stand that' *p*.¹⁸ In my present application of her account to the question of collective virtue and vice, I am applying these two different routes to joint commitment in respect of motive/skill, so that an individual group member may either reach joint commitment via her own motive/skill, or she may reach it by 'going along with' (minimally facilitating) the motive/skill in question. I think (ii) happens all the time, and typically takes the form of what I have called the 'practical identities' case. To take an example, imagine an appointments committee that is constituted by institutional fiat—constituted, for instance, by another committee that has the authority to do so. The committee may or may not have also constituted itself as a plural subject yet, but in any case some number among them have been in touch with each other, jointly committing to conduct themselves in a manner that promotes equal opportunities. (Perhaps they think insufficient efforts have been made in recent appointments.) When the whole committee meets up, the other members go along with the idea for broadly the right reasons, even while as individuals they don't particularly care that all much, and might not have noticed or bothered if others had

¹⁷ An extra anti-summativist twist might be that there can be cases in which it is impossible for any individual alone to achieve the requisite reliability—perhaps because it would require too much multiple expertise to do so. Our example of the diligent and thorough science research team could be like this, if their experiments called on sufficiently diverse kinds of high-level expertise.

¹⁸ '[I]n general, participating in *believing that p as a body* does not require *personally believing that p*. ... It is not even a *necessary condition* of our believing that *p* that each or most of us believe that *p* personally' (Gilbert, 'Remarks on Collective Belief', *Socializing Epistemology: The Social Dimensions of Knowledge*, ed. Frederick F. Schmitt (Lanham, Maryland: Rowman & Littlefield, 1994); 251).

not led the way—these are the passengers in the now committee-wide joint commitment, and they are the ones for whom a gap between two of their practical identities (private individual; member of the committee) has opened up.

The second way in which a group may achieve a virtue without all its members individually possessing the good motive/skill is if all of them commit as passengers only—they *all* merely ‘go along with’ the motive/skill of conducting themselves in a manner that promotes equal opportunities. If we imagine again an appointments committee, constituted within an institution that has explicit, comprehensive and effective procedures in place to promote equal opportunities, then we might say that the commitment to promoting equality of opportunity automatically comes with the job of serving on the committee. In such a scenario we can imagine that the entire committee—in the right spirit, if a little sceptical regarding laborious top-down directives—merely ‘goes along with’ the good motive/skill. In such a case it would be churlish to withhold the title of virtue from the committee, provided of course that it does indeed reliably achieve the good end, namely, equality of opportunity.

What these two possibilities reveal is that even within the bounds of the Gilbertian model of joint commitment, the unanimity requirement is more differentiated than it may at first seem. Further, it can be applied in a way that exhibits a good fit with our intuitions about how social groups function and the different ways they may come to display collective virtue and vice. The two different routes to joint commitment also prompt a useful ethical distinction in relation to individual credit, which fits well with virtue theory generally. Members of a sub-

group who bring their own personal good motive or skill to the larger group make a stronger contribution to the collective virtue than do passengers who merely go along with those who have taken the lead. Members of the sub-group therefore deserve more credit than the passengers, who contribute only weakly to the collective virtue.

Before moving on to the question of institutional virtue we should reflect on a third possibility. What about group members who are not even proper passengers (they do not even ‘go along with’ the motive, and so are no part of the joint commitment)? They are mere *stowaways* in the journey of the group. Such members are not in fact members of the plural subject that possesses the virtue, despite their being members of the appointments committee (to continue with our example). Their presence entails that, strictly speaking, the committee *per se* does not display the virtue, but only those members who are party to the relevant joint commitment (i.e. all bar stowaways).

I mention this possibility of stowaways partly because it depicts a familiar scenario, but also because it is worth noting that my account entails that the committee as a whole in such a case would not possess the virtue. This fact acquires some ethical significance if we imagine instead a case of collective vice, where there is a stowaway on board who is not party to the joint commitment to the vicious motive or deficit of skill. She is not a member of the vicious body, and so not party to the collective blame. This seems appropriate, not only because she has nothing to do with the vice, but also because there can be many social contexts of group activity where it is a real ethical achievement to resist colluding with the collective vice that may be a prevalent and powerful force in the environment. It is an ethical

achievement to refuse to go along with a joint commitment to a racist ‘canteen culture’ (of superficially friendly racist nicknames and jokes, for instance) in a working environment heavy with peer pressure, in which many of the white majority do at least ‘go along with’ that racist culture and ostracize those who don’t. This pressure to go along with the racist culture is a pressure that makes itself felt not only on the white majority but also on black members of the group. Hence being a stowaway can sometimes be the best a decent person can do in a bad social environment, and even that may require some significant courage and sacrifice. Further, for someone who had been jointly committed to the racist canteen culture but who wants now to withdraw, the appropriateness of rebuke from other parties to the commitment reminds one of the internally generated coercive power that such group phenomena possess.¹⁹

The burden of this section has been to elaborate and defend a differentiated version of Gilbert-styled joint commitment as our model of collective virtue. I now turn to confront the question of institutional virtue and vice.

3. Virtuous and Vicious Institutions?

So far the possible subjects of virtue have been individuals and collectives. For obvious reasons it has been the latter that has taken up our philosophical efforts. But we are now in a position to explain in what sense institutions might be said literally to

¹⁹ Gilbert notes this pressure to conform in connection with group belief; see for instance ‘Collective Epistemology’, *Episteme* Vol. 1, No. 2 (October 2004); 102.

possess virtues and vices. If someone describes, for instance, a certain child protection service as negligent, what is being said? Recall the three aspects of institutionalized groups that I listed at the outset. One thing that might be intended in such a statement is (1), that individuals working in the service are negligent in how they realize the procedural structures. This might apply to individual managers at whom the buck stops, or to an individual social worker with soul responsibility for a family containing a child that has not been adequately protected. Another is (2), that collectives in the service are negligent in how they realize those procedural structures. This might apply to a given team of social workers with joint responsibility for a practice of information sharing that has not been properly carried out. Yet another might be (3), that the procedural structures themselves are in some way ‘negligent’. This might be so if there is no effective mechanism for passing on cases when someone is off sick or leaves the service. I have spent the last two sections vindicating (2), taking (1) as read, and so I now attend to (3).

Prima facie it might look as though we can attribute virtues and vices to institutional structures and procedures. Certainly we might naturally describe a child protection service as negligent if it lacked a mechanism for passing on cases from one social worker to another. Nonetheless, it would be too much of a stretch to take this literally as an ascription of vice to the institutional structure itself, considered in isolation. We cannot literally apply concepts of virtue and vice to institutions considered purely procedurally, for they do not have any of the features of the normal subjects of virtue and vice: they are not agents, they have no will, and so they cannot accrue any credit or discredit of the praise and blame variety. I contend, therefore, that it is only in combination with the individuals and groups whose work realizes an

institution's procedures at any given time that the institution may literally be said to exhibit virtues and vices. It is people acting, whether as individuals or teams, that permits the institution as a whole to be described in terms of virtue and vice. Structures and procedures *in themselves* cannot—they form the skeleton of the institution, which must be given flesh by the people whose activity animates it. A well-structured skeleton is of course vital: clear and transparent procedures for declaring interests, making complaints, contesting unequal treatment, ensuring checks and balances, and so on, are of the utmost importance. But that is precisely because they encourage virtuous behaviour on the part of the individuals and groups whose activity makes whole the institution at any given time. To encourage virtue and discourage vice must surely be a fundamental criterion of good design in an institutional structure. To answer my title question explicitly, then, institutions can possess virtues and vice; and they do so only when their structures are brought to life by the individuals and collectives whose activity puts flesh on the institutional bone.²⁰

Quite which virtues it is desirable for different sorts of institution to display will no doubt vary, but there may be some very basic institutional virtues that are classed as basic because they are of fundamental importance to the legitimacy of the polity in general. I would like to end by suggesting that institutional virtues of epistemic justice, most notably the virtue of testimonial justice, are basic in this sense.²¹ An agent who is testimonially just reliably neutralizes the impact of prejudice

²⁰ It is easy to see that a well designed institution would be virtuous so long as the people whose activity realizes its structure behave virtuously; one can only speculate how much scope there is for a badly designed institution to still be virtuous, against the odds, thanks to specially virtuous conduct on the part of the people who realize it.

²¹ For an account of testimonial justice considered as an individual virtue, see my *Epistemic Injustice: Power and the Ethics of Knowing* (OUP, 2007).

in her judgements of speakers' credibility. A collective (such as a complaints committee, an appointments panel, or an investigating team of police officers) will possess this virtue if its members jointly commit to neutralizing prejudice in its judgements of speakers' credibility. How is this virtue fundamental to the legitimacy of our institutions? The claim might be made out by reference to various models of freedom and democracy, but let me try to make the point here by reference to the generic importance of *contestability*—a citizen's power to contest wrongful treatment of various sorts, notably wrongful treatment at the hands of institutional bodies, but also private individuals, as is the case when someone contests a theft or assault by reporting it to the police. Although this is a generic notion, it will be helpful to address the issue in a specific formulation. I will illustrate the point by way of Philip Pettit's republican formulation, because it involves a particularly explicit statement of the conditions of contestability.²²

In Pettit's account of freedom as non-domination, he models non-domination as a kind of security against arbitrary interference. Being required to pay one's taxes, or obey the law, is non-arbitrary interference. But if you are married to a man in a legal system that technically gives him the right to rape you, because it does not recognize the category of rape within marriage, then you lack freedom. For even if he is a good man who would never hurt you, the point is that were he to do so, you would not be able to contest it. And so you are dominated. It is of course an inevitable fact of life that people will sometimes be fired without due reason, receive unequal

²² See Philip Pettit, *Republicanism: A Theory of Freedom and Government* (Oxford: Clarendon Press, 1997). I thank Al Coffee for directing my attention to the connection between Pettit's contestability condition and my conception of testimonial justice.

treatment, suffer assault, and so on, and so what is needed to cancel the arbitrariness of these forms of interference is the real ongoing opportunity to contest one's treatment after the fact: 'what is required for non-arbitrariness in the exercise of a certain power is not actual consent to that sort of power but the permanent possibility of effectively contesting it'.²³

What is crucial in the present connection is that it is an explicit requirement of contestability that there be a *channel or voice available – either one can speak oneself, or there is some body that can speak with credibility for you (a representative)*. It follows from this condition that a contestee's word must be heard without prejudice. The hearer—and it will often be a group rather than an individual—must therefore possess the virtue of testimonial justice. A complaints panel, or court, or police force that lacks that virtue, allowing the credibility it gives to certain speakers to be deflated by prejudice, effectively prevents the speaker from contesting his treatment. And so he is dominated. It is clear from the Macpherson report that the way the investigating officers perceived and heard the word of the chief witness and fellow victim of the crime, Stephen Lawrence's friend Duwayne Brooks, signally lacked the virtue of testimonial justice. At paragraph 5.11 the report says:

the officers failed to concentrate upon Mr Brooks and to follow up energetically the information which he gave them. Nobody suggested that he should be used in searches of the area, although he knew where the assailants had last been seen. Nobody appears properly to have tried to calm him, or to

²³ Pettit, *op.cit.*; 63.

accept that what he said was true. To that must be added the failure of Inspector Steven Groves, the only senior officer present before the ambulance came, to try to find out from Mr Brooks what had happened.

Generalizing from examples of testimonial injustice such as this, we can see that a society whose institutions lack the virtue of testimonial justice is a society in which there is no real contestability for those who tend to be subject to prejudice. No real contestability, therefore, for those citizens who are most likely to need to contest their treatment. In such a society, interference in many citizens' lives remains arbitrary; and to that extent the society is not free. Such are the issues that turn on the presence or absence of institutional virtue.

4. Conclusion

In summary, I hope to have explained how institutions can, quite literally, possess virtues and vices; and I have argued, in particular, that the phenomenon of non-individual-based institutional vice needs clarification in philosophy and in public consciousness. I have focussed on the example of institutional racism in my efforts to make this point, but the issue is of course general. Indeed there are virtues—and I have argued that testimonial justice is one such—whose systematic absence from institutional bodies destroys many citizens' ability to contest wrongful treatment, and so belies a society's claim to freedom.²⁴

²⁴ I presented earlier versions of this paper at the Intellectual Virtue and Vice conference held at California State University, Fullerton in June 2008; at the Collective Epistemology workshop at the University of Basel, Switzerland, in

Miranda Fricker
Birkbeck, University of London

October 2008; and at our departmental seminar at Birkbeck in February 2009. I am very grateful to all those concerned for constructive comments and questions, and in particular to Michael Garnett, Margaret Gilbert, Anil Gomes, Jennifer Hornsby, Thomas Hurka, Miriam Solomon and Deborah Tollefsen for helpful informal discussions on those occasions.