

I imagine that we'll talk mainly about the body of this paper. I'd be grateful for reactions to the supporting material that I've put into two appendixes, but you'll have seen the central line of argument to its end if you stop reading at p. 12.

A. J.

Getting people to do things

A. J. Julius, November 9 2009

Some ways of getting people to do things are wrong. Early sections of this paper characterize that wrongness by defending this principle:¹

- (1) You should not (do y and believe that your y'ing will lead me to x and that this fact is a reason to y) unless, for some reasons R that hold independently of your y'ing, your y'ing helps me to [do x as a result of my belief that I should x in virtue of R].

Later sections use the principle to explain the point, scope, and content of certain political standards.

I

I believe that (1) explains the wrongness of central examples of wrongful coercion. They are wrong, I think, because they violate (1).

For examples take

- (2) threatening to shoot a person unless she hands you her money
threatening to fire a worker if she attends a union meeting
threatening to expose a colleague's adultery unless she votes for you

When you develop these episodes in the usual way, you suppose about each of them that the speaker takes as a reason for making her threat that it leads the listener to do something. And in each of them it's supposed that the listener, if she does that thing, will do it for reasons that depend for their existence on the threatener's having made her threat.

By contrast notice that typical warnings are OK as far as (1)'s concerned. If I know that in my rage at your holding onto the money I will hurt you, and if I tell you this beforehand to save you a beating, you have sufficient reason, independent of the warning, to do what I'm trying to get you to do. The hypothesis that wrongfully coercive threats are wrong because they violate (1) also explains the appearance that typical warnings are not wrong in the same kind of way.

Of course (1) is not the only principle that accounts for the wrongs of (2). In condemning the mugger's threat to kill the muggee you might instead appeal to

- (3) If your z'ing is impermissible in virtue of its consequences or its significance for me, and if your y'ing gives me reason to expect that you'll z, then your y'ing is impermissible.

Threats to do independently impermissible things are perhaps themselves impermissible.

But I doubt that (3) locates the central objection to objectionable coercion. It seems to me that a typical (2)-style threat—a threat to expose your adultery, say—is morally worse if and because it's part of my plan to steer you into some course of action like voting for me than if I do it on a whim. Moreover I'm pretty sure that a threat to do something permissible can count as wrongfully coercive. I coerce you

and I wrong you when I threaten to sell a couch of mine that holds fond memories for you unless you do my chores.

I'm no more optimistic about views that trace coercion's wrongness to the wrongness of making a person worse off, or of making her unfree, by depriving her of options for action. By threatening to spit on you if you take a seat on the bus I deprive you of the option of sitting there unspat upon. But I subtract that same option—subtracting from your wellbeing and from your freedom whatever was subtracted by the threat to spit—when I sit down there myself. Only the first of these option removals seems wrong. So these views threaten to overproduce judgments of wrongness. Elsewhere they tend to underproduce them, as when they fail to condemn my act of threatening to leave money to a worthwhile charity that you disapprove of unless you polish my grave. That threat carries no substantial cost to any person's wellbeing or freedom, but it is wrong.²

Here's an argument for (1), then. If it seems to you that the acts in (2) and in (4) threatening to sell my couch unless you wash my clothes threatening to give to a charity unless you polish my grave are wrong, and if you think that they have in common some feature that makes each of them wrong, and if you think that these are not true of the examples in (5) taking a seat on the bus before you sit there threatening on a whim to expose your adultery then you should take as evidence for (1) the fact that, of the principles I've mentioned so far, only (1) agrees with you about all of that: the acts and attitudes in (2) and in (4) violate that principle while the acts and attitudes in (5) are untouched by it.

This drive-by argument is not too effective. I'm not too confident in the judgments I've just rattled off. I know that many people don't share them. And of course I've left many alternative explanations of them still standing.

II

A more powerful argument will have to explain the explanation—explain why violations of (1) are wrong. If there are some things that are just the case, unexplainedly and inexplicably the case, (1) is not among them.

The courses of action that violate (1) are courses of directing the action of other persons. This principle takes them to be wrong because of the bad way in which they direct it. To defend the principle I'll have to explain how action-directingness can matter to wrongness. And this looks unpromising. I discuss one difficulty—general anguish about how motives can matter to wrongness—in the first appendix.

The other problem is that a good explanation is hard to find. You will hear that coercing a person marks a failure to recognize her agency or to respect her autonomy. Probably something like that is true. Left there, as an explanation of coercion's wrongness, it fails to satisfy. If I threaten to tease your sister unless you jump through a hoop, and if you then jump through it, you are exercising your reason, jumping because you see that my threat to tease counts in favor of doing

that. I make my threat precisely *because* I know that you're a reasoner and that I can count on your reasoning toward the act that suits me. How here do I fall short of recognizing or respecting your agency or autonomy?

That is not a rhetorical question. We need to find out what it is to take heed of a person's agency or to respect her autonomy such that these attitudes are inconsistent with coercion or with violations of (1). I doubt that this can be discovered in a direct interrogation of agency, autonomy, recognition, respect. The strategy I like better is to explain the moral onus on action-directing action with the help of considerations *outside* this circle and then to bring those considerations in as lending content to agency, autonomy, recognition, respect. I'd like to solve for those values subject to the constraint that they're what's at stake in getting other people to do things. You can spell "respect," now find out what it means to me.

III

I'll argue here that following (1) is required. That violating it is *wrong* in particular will have to come later.³

A

Forget other people. Think about the intrapersonal ethics of your own government of what you do. You should, I think, act for reasons. Or, better to say, you should not act not for reasons. You face a rational requirement like

(6) You should not (do x and have it be false of you that, for some reasons R, you [do x as the result of your belief that you should x in virtue of R]).

This principle is false, of course. It's too strong. You can, without violating any plausible rational requirement, do certain things routinely or idly or on a whim and so without believing that you have good reason to do them. And it's not clear how best to weaken (6) so as to formulate these dispensations. But my interest allows me to set them aside, for a reason that I'll explain in footnote 4.

Some people reject (6) all down the line. They deny that any such revision can make it plausible that you face a genuinely normative requirement to act for reasons. You should do, if anything, at least one of the things that you have sufficient reason to do, they think. And, they think, you should believe it true, of those things, that you have sufficient reason to do them. But these requirements to do the reason-supported acts, and these requirements to believe that the reasons support the acts, exhaust the requirements that bear on those acts and beliefs. Beyond those simple act- or belief-governing principles, there is no compound act-and-belief-regulating requirement demanding of you, as (6) demands, that you not act not for reasons.

B

The fact that *you* should do things for reasons rationally constrains the attitudes that *I* may take toward your action and the acts by which *I* may affect it. A requirement like (6) is *public* in the respect that, when read as referring to you, it bears not only on what you do and how you reason but also on my action and practical thought.

Here's a simple way of making this interpersonal connection. Too simple. You might think only that I should act in a way that allows you to satisfy (6). That stricture would make some difference. It would tell me not to put science-fiction stuff into your head that causes you to jump without your believing that you have reason to jump. But it wouldn't furnish anything like (1). By threatening to break your fish tank I can bring it about that you cook my breakfast. But you'd be cooking it as the result of your belief that, thanks to my threat, you have good reason to cook. And so the threat's compatible with your satisfying (6).

The interpersonal upshot I have in mind is trickier. To explain it I need to take another look at intrapersonal problems.

C

You and your friend are unhappy together. You see that you should leave and that, once gone, you should keep your distance. Behind a radiator in the friend's apartment you hide the expensive fur hat that you had borrowed from your employer. Your secretion of the hat makes it the case that you'll have good sheerly hat-grounded reason to return. You leave it there so that it will bring you back.

You shouldn't do something without believing that you have good reason to do it. And you shouldn't try to *bring it about that you do something* except where you believe that you'll have good reason to do *that something* in virtue of reasons that hold independently of what you now do to bring yourself to do it.

Here's a stab at the second thought:

(7) You should not (do y, believe that your y'ing will lead you to x and that this fact is a reason to y) unless you have, independently of your y'ing, good reason to x.

But this principle is too weak, as the next idiotic example will make plain.⁴

D

You're back with your friend. I forgot to mention that your mother lives in the same apartment complex. She's very sick. The fact that by returning to the complex you'll bring her some comfort is a decisive hat-independent reason to return. But you have issues with *her*, too. Those might keep you from carrying this out. So maybe you should leave the hat after all?

I think that it would be OK of you to leave the hat and to take as a reason for leaving it that you'll be led to do something that comforts your mother. But to leave the hat while taking as a reason for leaving it that you'll be led to resume your unhappy connection with your friend is not OK.

(7) is blind to this difference between two ways of stashing a hat. A principle that picks it up is

(8) You should not (do y and believe that your y'ing will lead you to x and that this fact is a reason to y) unless, for some reasons R that hold independently of your y'ing, your y'ing helps you to [do x as a result of your belief that you should x in virtue of R].

This rules out hiding the hat on Monday so that you'll reconnect with the friend on Wednesday. Produced in that fashion your return to the apartment will count as the result of your judging on Wednesday that you should retrieve the hat and of your

wanting on Monday to see your friend again. It won't be the result of any day's recognition of the mother reason that you judge to suffice for returning and that's independent of the hat's location. By contrast (8) rules *in* your hiding the hat on Monday so that you'll see your mother on Wednesday. For there your return will be the result not only of your Wednesday recognition of a hat reason but also of your Monday recognition of a mother reason.

E

The bad news is that I want to talk about another principle. The good news is that, if I'm right, it's the only principle you're going to need in this area.

Consider

(9) You should not (do *y*, do *y* as the result of an attitude of yours in favor of A's *x*'ing, and have it be false of you that, for some reasons *R* that hold independently of your *y*'ing, your *y*'ing helps A to [do *x* as the result of A's belief that A should *x* in virtue of reasons *R*]).

(8) is the specialization of (9) that you get by substituting yourself for A and by substituting, for an attitude in favor of A's *x*'ing, your belief that your *y*'ing will lead to your *x*'ing and that this is a reason to *x*. Though a little harder to make out, (6) is also very nearly a special case of (9). Just let A be you and let *y* be *x*. Then you have

(6') You should not (do *x*, *do x as the result of an attitude of yours in favor of your x*'ing, and have it be false of you that, for some reasons *R* that hold independently of your *y*'ing, your *y*'ing helps you to [do *x* as the result of your belief that you should *x* in virtue of *R*]).

This differs from (6) only in the italicized phrases. For "your *y*'ing helps you to" substitute "you". The insistence on reasons "that hold independently of your *y*'ing" makes (6') more plausible than (6) by insisting, plausibly enough, that the reasons for which you act be facts that hold independently of your act. And the first phrase makes (6') more plausible than (6) by restricting its attention to stuff you do as a result of attitudes in favor of that stuff. Presumably your knocking over the cookie jar despite your lacking any belief that you should knock it over violates a rational requirement only where the upsetting of the jar figures in the content of the attitudes that lead you to do it.

If it seems to you that (6') is true, and that (8) is true, and that those two principles are of a piece insofar as they require that your production of your acts be governed by your recognition of reasons for those acts that hold independently of their production, then you should mark in (9)'s favor the fact that it's an exact general requirement to that effect and that it subsumes the principles (6') and (8).

F

I've been talking about intrapersonal instances of (9) for which A is you. But why suppose that (9) is true only where A is you? That is not a rhetorical question either. But until I can answer it, I'm ready to believe that the full interpersonal principle—(9) for any value of A—is true.

I mean to appeal to a thesis of publicity:

(10) If A should have some attitude toward A's *x*'ing, and if B can also have that attitude toward A's *x*'ing, and if nothing in sight counts against

B's having that attitude toward A's x'ing, then B should also have that attitude toward A's x'ing.

Consider the attitude *trying to bring it about that A does x, if at all, only as the result of A's recognition of reasons that hold independently of your bringing this act about*. I've argued that you should take the version of this attitude for which you're A. But you can also have the version of it for which I'm A. And the publicity principle (10) says that, in the absence of considerations against your having it, you should have it. So there's a case for (1).

IV

A

Coercion is not the only thing that (1) condemns. Force, deceptive manipulation, and plays upon persons' emotions are also often forbidden by it. So are certain offers.

This brings an objection. Most offers are not, after all, wrong like coercive threats. When from across the cafeteria table I tell you "I'll give you my chips if you give me your pretzels" I'm leading you to do something—surrender a salty snack—that, as it may be, you don't have sufficient reason to do independently of my offer. Apparently this violates (1). But it is no less apparently irreproachable.

B

Though it might be that you don't have sufficient reason to give me your pretzels, it might also be that, because pretzels are good for me and chips good for you,

(11) We should do (I give you the chips, you give me the pretzels).

It might be that you and I are subject to a *joint requirement* that we perform that set of acts—that (11) is true of us, and that there is no conjunction of individual requirements of the form *I should do this, you should do that* such that (11) consists in or holds in virtue of the conjunction. I argue in the second appendix that there are some joint requirements.

If (11) holds at all, it holds independently of my offer. And my offer aids in your playing your part in the pair it requires. You can't do your part unless I will do mine, without which there's no pair for yours to form part of. Letting you know that I'm ready to do mine allows you to opt for doing yours as a way of fulfilling (11). My offer helps you to come to do your part as the result of your accepting this offer-independent requirement that you and I together face.

I conclude that offers are OK by (1) where they can be understood as facilitating the other person's response to a joint requirement of the pair of acts that carry the offer out.

C

Not every offer is susceptible of vindication by this proposal. It is false that

(12) We should do (you work in my sweatshop to produce \$20 an hour, I give you \$1 an hour).

So my offer "I'll give you \$1 an hour if you work in my sweatshop" doesn't facilitate your response to an offer-independent joint requirement of carrying it out.

This observation is maybe the germ of a theory of exploitation. Exploitative offers, I imagine, are the ones that can't be justified as facilitating their recipients' response to a joint requirement because the relevant acts are too lopsided, too asymmetrical in their reason-giving aspects, to be jointly required. Though they're not coercive, these offers draw a moral objection from the same source as condemns wrongful coercion: from (1), the morality of getting others to do things.

With this story about offers tacked up as a model, I turn to some other political incidents of that morality.

V

An idea that is much in the air.

(13) For some standard *S*, the fact that an institution satisfies *S* makes it the case that the acts of coercion that help to cause or constitute the institution are not wrong. We should bring it about that the institution satisfies *S* because this makes the relevant coercion not wrong. A society's being just consists in the fact that its institutions satisfy *S* or in the fact that, because they satisfy *S*, the coercive acts that cause or constitute them are not wrong.

The *point* of justice, seen from this view, is to purge our institutions of wrongful coercion. A patch of social life can be just or unjust only if it's regulated by coercive institutions that raise the one moral problem that only justice can solve. So the *scope* of justice coincides with the scope of those institutions. The *content* of justice's standards is determined by such considerations as justify coercion, and they owe their *urgency* to the requirement that we not wrong people when we coerce them.

These proposals are hard to get down. Here are a few difficulties.

Point and scope Why aim for a Rawlsian well-ordered society where people obey the liberal-egalitarian laws in the conviction that they're correct rather than for a nightmarishly inequality-ridden Lockean association of independent producers who obey the property laws because they're convinced Lockean? What reason does *justice* give you to aim for the first society? Each association is uncoerced. So the two are tied in respect of avoiding wrongful coercion. To break the tie you might take on

(14) A social arrangement's being just consists in the fact (a) that if coercion were involved in causing or constituting it, this coercion would be permissible; and (b) that it has some further morally valuable features.

But then it's tempting to find the whole nature of justice in (b). Justice is having the (b)-features, whatever they are. Having them justifies any coercion that arises. But that justification of coercion is no part of their justice.

Many people find it hard to believe in any case that standards of justice come into force only where people are trying to coerce one another. They think that these standards already recommend to people who've never before interacted that they aim for the strong patterns of allocation or interaction that justice demands.

Content and urgency A principle like (1) leaves no explicit room for morally decent coercion. If coercion is always wrong, there's no hope of working out the content of justice standards by asking what can make it OK. On the other hand it seems that many acts of coercion are not wrong all things considered. But these appear to owe their permission to their promotion of or respect for urgent moral interests whose existence and force are independent of any moral constraints on coercion. The interest in protecting people from being killed, for example, justifies a coercive threat to imprison killers. Because these urgent independent interests are urgent and independent, the standards through which we honor or pursue them in our social life won't depend for their content or urgency on the avoidance of wrongful coercion. But then it's looking unlikely that a concern with coercion will impart a distinctive content or urgency to our political standards after all.

These complaints sound right to me, and they incline me against the coercion view. I'll now describe an alternative, centered on (1), that stands clear of these problems even as it joins the coercion view in claiming that moral requirements on action-directing action ground some special political standards.

VI

Though (1) covers other sins, let's hold the focus on coercion for a moment. What if anything can make it the case that coercion is not wrong?

A

A boring but credible suggestion. Often enough, declining to secure some urgent moral interest by coercive means draws an objection even stronger than (1) raises against the coercion. Then violating (1) is not wrong considering all things.

Perhaps the moral badness of severe poverty is enough to license the coercive enforcement of taxes to help poor people. But the moral interest in furnishing poor people with the things they lack has got to be strong, quite strong, if it's to overcome (1). Presumably that weighty interest already recommends with all its coercion-independent weight that political communities act to end poverty even before coercive projects of tax collection cross their minds.

This is the content problem that I described in the last section. The problem goes away when a second, maybe more interesting justification of coercion comes in.

B

Every dweller in our low town has the same interest in a levee's being built to hold back the floods. In virtue of those interests we face this joint requirement:

(15) We should do (each pitches in with labor and materials to make a levee that protects us all).

Suppose we're uncertain whether every person will recognize or act on (15). Then we might announce that, if anyone holds out, we'll pump water into her house, taking as part of our reason for this announcement that, if a person doesn't in the end recognize (15), the announcement will lead her to contribute.

I claim that this announcement is reconciled with (1) by the consideration that, if I do in the end recognize (15), the announcement will have helped me to contribute

as a result of my belief in (15). Where the announcement turns out *ex post* to be coercive, it's justified by the *ex ante* prospect that this same announcement helps to make it the case that I contribute because I accept (15).

The announcement does this work by assuring me that the others will contribute. When I look forward to seeing them heave their sandbags onto the wall, I also know that by showing up with my own bag I can play my part of the set mentioned in (15).

But the form of this assurance is quite special. I don't claim only that the utterance of this announcement addressed to the others assures me that they'll contribute. For this wouldn't show that the addressing of it *to me* aids my own response to (15). And that's what required if the announcement I hear is not to violate (1).

My thought is instead that the threat's being addressed to me helps me to contribute on the basis of my acceptance of (15), by helping to make it morally OK to address the announcement to every other person in a way that assures me that she will do her part. It's a criss-cross justification:

(16) For every person *i*,

(A(i)) We address the announcement to *i* in part so that, if she doesn't in the end recognize (15), she nonetheless contributes;

and A(i) is justified by the fact

(B(i)) We address the announcement to *i* in part so that, if she does in the end recognize (15), she can act on it;

and B(i) is justified by the fact

(C(i)) For every *j* other than *i*, A(*j*) holds, assuring *i* that *j* will contribute, and B(*j*) and C(*j*) hold, justifying A(*j*).

Here it's true of every person, every node of this totally connected justificatory graph, that addressing the announcement to that person helps her to respond to (15) by ensuring that the others can be led to do what (15) asks of them without being themselves wronged in the eyes of (1).

VII

A

For any group of us, there are things we should do. We have several ways of getting the members of our group to do these jointly required things. We can make potentially coercive announcements like the ones I just considered. Each can commit herself to following centrally chosen rules and then try with the others to establish rules directing each to play her part of the jointly required sets. We can try to persuade one another that each should play her part because the set is jointly required while assuring each that the others will play theirs by effecting their persuasion. Beyond these direct means of arranging our acts into the required set by invoking that set as part of the content of our coercive announcements, our rules, or our arguments, we can also use coercion, rules, and arguments to bring action about whose own effects will in turn lead people to do the jointly required acts. Though each human's next move is unpredictable—that's why we love the human!—the central tendency of the action of many humans is somewhat more predictable. We can make it the indirect target of our coercion, our rules, and our arguments.

One of the things we should do, I think, is to induce a mixture of coercion, jointly adopted rules, argument, and the indirect manipulation of people's action by manipulation of the foreseeable consequences of coercion, rule, and argument—a mixture of those that will lead us all to do the various other things we're jointly required to do.

This action-directing ensemble has two virtues. First, there's a bunch of stuff that we should do, and by aiming for this ensemble we can bring ourselves to do all of that. But, second, we can uphold (1) in our attempts to get people to do things if we aim for this ensemble. This is unproblematic where the argumentative and the rule-making components are concerned: rational persuasion and the fine-tuning of rules that each undertakes to obey as a way of doing her part are morally decent ways of directing people's action. But the potentially coercive and the indirectly manipulative components are also eligible for approval by (1). For the ex post coercion or manipulation of a person might be justified in the "criss-cross" fashion of VI.B; it might be justified as helping that person [to do her part because she accepts that the set is required.]

B

I'm not quite ready to identify a society's being just with its having this ensemble in place. Having these *institutions* in place, I might as well say. But consider what such an equation would make of the problems of point, scope, content, and urgency.

Point One point of aiming for the just ensemble is strictly negative. We can avoid wrongfully misdirecting other persons' action if we direct their action only by aiming for the just-making institutions. But there is a second, positive rationale for that program. Between aiming for a just uncoerced association, aiming for one that's not just, and refraining from association altogether, we should go for the just one because we're jointly required to do the things we're led to do under it.

Scope Every person shares with all other persons near, far, past, present, and future—no matter whether and how she may now be interacting with them—a reason to aim for the just society so that we do things that we should do. In this cosmopolitan conclusion the proposal agrees with divided-world luck-egalitarians and other desert-island cosmic-justice types. But unlike them the proposal can't make out a presently effective individual duty to aim for a just world society. Each of us should now expect most other people to fail to participate in the institution-making that's jointly required of us on a world scale. So none of us has in virtue of that requirement alone an individually effective reason to get it started.

For some *content* and *urgency*, at last, take a look at three political standards.

VIII

A. *Equality*

I assume that, where people have similar interests in the common effects of all their actions, they face requirements to perform a set of acts that advances those interests in similar degrees. Not for deep egalitarian reasons, just on grounds of indifference. At high-enough levels of abstraction that similarity condition is

surely met: our interests in leading good lives, for example, are presumably indistinguishable. If that's right, we're jointly required to aim for sets of acts that, as it happens, distribute life's goods in a rough equality.

We can accept these requirements without supposing that the worst-off people can raise a stringent personal objection against every failure to meet them. An A who lives in the lowest ward has, I imagine, no stringent objection to B's, C's, and D's building a levee that protects only their higher ground, even if this construction falls short of the requirement (15).

But A should cry foul when B, C, and D try to impress her into building their too-low levee. For it's false that

(17) A, B, C, and D should do (each person pitches in to make a levee that protects B, C, and D).

Directing A to work on a levee can't be justified by something like (17), only by something like the egalitarian (15). B, C, and D wrong A by directing her to work on any but an egalitarian levee.

The boring justification of coercion held out for urgent moral interests strong enough to overturn (1)'s prohibition of wrongness. The interesting justification of coercion looks elsewhere. A joint requirement can serve not to override the principle (1) but to make certain ex post coercion OK by it. The requirement that plays this role, since it need not best (1), need not be grounded in reasons of any particular moral stringency. So where we affirm that (1) does ground stringent political standards their stringency can be the creature of (1) itself.

A's story makes a good example. Her loud complaint against being led to work for any but an egalitarian levee owes its severity to (1) and not to the antecedently mild requirement that the people build that kind of levee.

B. Liberty

The so-called interesting justification of getting people to do things runs on genuinely joint requirements. If each person's interests are affected only by her own toothbrushing, an apparent requirement like

(18) We should do (I brush my teeth, you brush your teeth)

reduces to a pair of individual ones like

(19) I should brush my teeth.

You should brush your teeth.

It's not joint. And there's nothing joint here to justify a coercion of your toothbrushing.

More generally where there's no interaction between the reason-giving features of what I do and the reason-giving features of what you do, there'll be no joint requirement that can make the extrarational manipulation of any one of us alright. So you have a kind of Harm Principle. Extrarational interference with my action gets a pass only where other people have the sort of stake in it that can make it part of something that all the people should do.

The view also contains, along with this Millian liberty, a Rawlsian one. To justify your directing a person's action requires at least some prospect that, starting where she starts, she will reach the judgment that she has reason to do what you're trying

to get her to do. If there's no sound deliberative route from her initial beliefs to that judgment, you fail that test. You can't credibly pretend to a facilitation of her response to the reasons she faces if there's no live chance of her coming to recognize them by reason alone.

C. Democracy

I've been talking about a project, justice, by which we can get ourselves to do things that we should do. If we do this right—if we honor liberty by directing our action in a way that takes account of our common status as persons whose acts are to be governed by their own beliefs about the independent reasons that support them—we'll have helped each person to act from her recognition that her action forms part of the equality-bent patterns in which we should act. Each will do what she will because it's part of what we should. I think that this fact will make it true of us that we act together, that acting together on large scales up to the scale of all humanity is valuable for its own sake, and that the value it realizes is democracy.

These thoughts have me thinking, of liberty, equality, democracy, and justice, that we can't have one without the others.

Appendix: motives, permissibility, and wrongness

“Suppose you were prime minister,” says Scanlon, relaying an example from Thomson,

and the commander of the air force described to you a planned air raid that would be expected to destroy a munitions plant and also kill a certain number of civilians, thereby probably undermining public support for the war. If he asked whether you thought this was morally permissible, you would not say, “Well, that depends on what your intentions would be in carrying it out. Would you be intending to kill the civilians, or would their deaths be merely an unintended but foreseeable (albeit beneficial) side effect of the destruction of the plant?” Holding fixed the actual consequences of the raid and what the parties have reason to believe these consequences to be, might an action be permissible if performed by an agent with one intention but impermissible if performed by an agent with a different strategy in mind? I agree with Thomson in finding this implausible.

Are they right?

The principle (1) implicates in coercion’s wrongness the coercer’s belief that her effect on the coerced’s action is a reason. I want to look at Scanlon’s arguments against motive-dependent morality, since they might make trouble for (1).⁵

I

By administering a strong dose of morphine to a patient who’s certain to die within hours you will both relieve her pain and cause her to die sooner, releasing yourself to an early start on your ski vacation. The reason you have to relieve the pain is, I’ll suppose, far stronger than the reason you have to prolong her life in pain for a few hours. It seems to me that

(21) You should give her the morphine.

and so that certainly

(22) It’s false that you should not give her the morphine.

On my way to those conclusions I didn’t stop to consider your beliefs about reasons, and in fact it seems to me that

(23) The truth of (22) is constitutively independent of your beliefs about reasons for giving the drug.

What room remains for common-sense morality’s condemnation of killing intended to let you leave for your vacation? Consider

(24) If you believe that the fact that giving morphine to a person will kill her and allow to leave for your vacation is a reason to give her the morphine, then your having that belief makes it the case that you should not give her the morphine.

If (24) is true, then (22) can’t be true of a person who has that belief. This denies the independence declared in (23).⁶

But (21) is not the only possible representation of a moral prohibition of intentional killing. Consider

(25) You should not (give morphine to a person and believe that the fact that giving morphine to her will kill her and allow you to leave for your vacation is a reason to give it to her).

In (25) the *should not* takes as its object, not a single act, but the (parenthetical) conjunction of an act and a belief. And it asserts that *should not* unconditionally on the actor's beliefs. So it's consistent with (22)'s being true of you whatever you believe. It can be at once true of you: that you should not (give the drug and believe that the vacation provides a reason); that you believe that the vacation provides a reason; and that it's not the case that you should not give the drug.

II

The move from (24) to (25) looks, well, like a move. A trick too cheap to solve any problems. Half the time it still looks that way to me. Now and then, and including *now*, I think there's something to it.

I think that, beyond the facts as to whether I should hold various beliefs and as to whether I should perform various acts individuated independently of the beliefs that accompany them, there are facts as to whether I should instantiate various act-and-belief combinations and as to whether my thought and action should follow various act-and-belief sequences. And, while we're at it, facts as to whether a single actor should perform sequences of acts. And, if you don't mind, facts as to whether the members of a population of actors should do various sets of their several acts. There are several *levels* to what we should do and how we should be. Above a ground level populated by single-act and single-attitude requirements, we face requirements to seek or avoid compounds and series of those ground-level things. Reduction is tempting, as usual. Maybe the compound requirements hold in virtue of or consist in conjunctions of the simple ones. Or maybe dependence goes in the other direction; it is rumored that, for every act that I should do, I should do it because it forms part of an extended acting-on-a-maxim that's required of me. But until you accept such a subordination, you have a competition: compound requirements vying with simple ones to regulate the objects that figure in both.

The problem of how I can answer at once to (21) in respect of my acts and to (25) in respect of my acts-and-beliefs is an example. Let me work through it a bit further. To organize the discussion I'll be asking whether (25) succumbs to versions or relatives of the objections that Scanlon raises against (24).

III

A

The charge against (24) might just be that my having an attitude can't make a normative requirement true of me. But that way of insisting on objectivity is not sane. The fact that I love my mother, which makes my presence valuable to her, can help to make it the case that I should stay home rather than join the resistance. Our requirements are not generally independent of our attitudes. The dependence

on attitudes that's supposed to underwrite (24) is suspect, if at all, only in some special way beyond its being such a dependence.

B

Dialogue's distracting. Consider a Thomsonian soliloquy:

Should I give the morphine? That depends on what I'd be taking as a reason for giving it. I should do it if I'd be taking pain relief as a reason.

But I shouldn't do it if I'd be taking an early vacation as a reason. So what would I be taking as a reason?

What's weird about the speech is that the speaker takes the answer to her deliberative question—should she do it?—to depend on her having or lacking attitudes—among others, her beliefs about reasons—that she's forming or reforming within the same course of thought through which she's trying to answer that question. Even the prince of Denmark was never so self-involved. Normal reasoning about what to do seems instead to take as its object normative facts that you take to hold independently of your performing the reasoning and of its results in your intentions and beliefs. The not so normal exceptions are important; I'll get to them soon. But this is how I experience a certain kind of normal thinking about whether to do a thing.

(25) is true to that experience. Unlike (24), (25) leaves open that your deliberation about which reasons count in favor of or against giving the morphine and about whether you should give it amount to the discovery of facts all of which hold independently of your deliberation and of the attitudes it produces or refines.

C

A principle that tries to formulate this kind of objectivity:

(26) The fact that you should *x* is constitutively independent of your beliefs about whether you should *x* and of your beliefs about such matters as help to determine whether you should *x*.

But this is still too much. My belief that the cause is just, because it makes me an effective combatant, can help to make it the case that I should join the resistance—a fact that also depends constitutively on the *fact* that the cause is just. But (26) would insulate the fact as to whether I should join even from this faultless dependence on my belief. It's not always a mistake to invoke, at some late stage in your reasoning about whether to do something, the fact that earlier in that reasoning you'd reached some conclusion about reasons for it.

D

In that example my reasoning about whether to fight breaks into my settling of three issues: whether the cause is just, whether I'd be good at advancing it, and whether I should join a just cause that I'd be good at advancing. The determinants of these issues are mutually independent; what makes the cause just, or what makes its justice and my high fighting capacity sufficient for my joining it, involves nothing of what makes me a good combatant. And vice versa. Their independence, I imagine, makes it okay for me to regard my belief about the first issue as fixed—as invariant with respect to the thinking about the second and third issues that lies ahead of me. The fact that I regard the second and third issues as open sits comfortably with my treating the first as settled and so with my placing my position on it out of the reach of my current reasoning.

This test is failed in other, objectionable instances of attitude dependence. Suppose I were to take the fact that I want to climb the mountain as helping to make it the case that I should climb the mountain. Then I'd be taking as fixed an attitude, my wanting to climb, that depends for its appropriateness on the very matters that determine whether I should climb: the danger, the view, my responsibilities to people on the ground. I shouldn't regard the question of whether I should climb as open without regarding those issues as open and so without treating as open the appropriateness of my wanting to climb. And so it'd be a mistake to treat that attitude as invariant with respect to my further deliberation about whether to climb.

E

Back to the doctor. Consider two deliberative issues: whether the ski weekend counts in favor of giving the drug and whether you should give the drug given that you believe that skiing is a reason. These issues are not independent in the sense just set out. Whether you should give the drug given that you take skiing as a reason depends on whether it is a reason. If you think both that skiing is a reason for drug giving and that your having this belief counts against your giving the drug, then you have some more thinking to do: you ought to reopen the whole question of whether skiing counts in favor of drug giving. It's a mistake to press ahead with your deliberation about the second issue while taking as a constant of that deliberation the fact that you've answered the first as you have.

F

Suppose that taking your belief as fixed in this way is indeed a mistake. I want to conclude that your having the belief does not help to determine whether it's the case that you should give the drug. It does not count against your giving it. For if some fact helps to make it the case that you shouldn't do some act, it can't be that your counting it as a reason against the act is a mistake. The circumstances that set the boundaries of our practical problems and that ground the requirements that apply to them can't include the fact that I have some attitude unless I can regard that fact as constant and still deliberate well.

I propose that (24) goes wrong not by violating objectivity-as-independence-from-attitudes but by putting reasoners in the bad position of treating as fixed attitudes that they ought not to take as fixed—by trespassing against a principle like

(27) If the appropriateness of holding an attitude *A* at *t* depends on matters that you should be attempting to settle in your thinking, at *t* and beyond, about whether to *x* at *t* + 1, then the fact that you have *A* at *t* can't help to make it the case at *t* that you should *x* at *t* + 1.

For its part (25) passes this test that by eschewing any recommendation of acts conditional on beliefs.

IV

I've offered a way of understanding the apparent problem of attitude-dependence on which it turns out to make no real trouble for (25). In deciding what to make of this, you might also consider that (25) is not alone in raising this problem by governing acts together with attitudes.

A

Another Thomson story. Because of a freak electrical disturbance, flipping a light switch in your apartment will cause a lightning flash that kills your neighbor. You don't know this. Your earlier experience with the switch makes you justified, let's assume, in believing that turning it on is harmless.

Thomson claims that it's impermissible to turn on the switch. This is presumably in virtue of something like

(28) If flipping the switch will hurt someone, you should not flip it

Scanlon denies this. He defends instead a principle like

(29) If the possible belief that flipping the switch will be harmless is not justified in your evidentiary situation, then you should not flip it.

and concludes that flipping the switch is permissible after all.

In defense of (28) Thomson imagines another dialogue. Suppose B has run the test and knows that switching is harmful.

A: Should I flip the switch?

B: If I tell you what I know, then it'll be true that you shouldn't. But so long as I don't, it's false that you shouldn't.

Scanlon says that what's wrong with B's speech is that it presupposes that A's coming to believe that the act is harmless will change its permissibility. And of course that's not true on Scanlon's (29) since permissibility depends not on whether A believes that the act is harmless but on the justifiedness of that possible belief.

B

But B's telling you that switching is harmful won't merely change what you believe. By changing your evidence it will also change the justifiedness of your belief that it's harmless. This conversation makes that belief unjustified and so by (29) it makes turning the switch impermissible. So what B says is true after all on Scanlon's view. But it still seems mistaken.

Suppose that, despite being justified in your belief that switching is harmless, you run the test anyway and find out that it's harmful. On Scanlon's view, by running the test, you make the act impermissible. But surely that's wrong. Surely you *discover* that it's impermissible.

There's evidence that Scanlon agrees. He says that when a person knows that switching is harmful "it is the fact that is known, not the fact of his knowing it, that he is to take as counting against his action." But this diagnosis seems excluded by Scanlon's (29). On (29) what counts against the act is neither the fact of harmfulness nor the fact that you believe it harmful but rather the fact that a possible belief in harmlessness is not justified. So in this last passage Scanlon depicts the person as acting on the basis of considerations apart from the ones that his own principle makes determining of permissibility.

C

Scanlon is surely *right* that Thomson's (28) is not the whole normative story. We seem to face another requirement. "One ought not to do what one sees, or should see, will cause serious harm to someone. One ought to take due care not to cause harm." This can be seen, Scanlon says, as follows:

If any principle that it's not reasonable to reject would require us to take C as a consideration counting strongly against an action, then, since it may not be immediately obvious whether C obtains or not, in considering any other principles it would not be reasonable to reject, [we should conclude that one of those other principles is that] we are required to be on the lookout for C, and to take reasonable steps to find out whether or not it obtains. [AJJ's interpolation]

I think the further principle might be this:

(30) You should not (flip the switch and fail to believe with justification that flipping the switch is harmless.)

D

If you start out to think about whether flipping the switch is harmful, you'll be guided not by (30) but by the evidence for and against harm and by the requirement that you form beliefs on the basis of evidence. And if you then conclude that flipping is harmful, you will use (28), not (30), to decide you against doing it.

Suppose on the other hand that your evidence warrants neither the conclusion that switching is harmful nor the conclusion that it's not. Then (28) fails to tell you what to do as you don't know which of your acts upholds it. But (30) is still in force. It offers two options: you can go on trying to find out whether switching is harmful so that, if you conclude that it's not, you can switch consistently with (30). Or you can simply not switch.

In the second case you'd be taking the fact that you don't have a justified belief in harmlessness to count in favor of not switching. Does this violate the test of section IV? No. For again you can decompose your question into two issues. Is switching harmful? And is switching supported or opposed by any facts about what you believe about harm? In taking the second question to be open you need not hold the first open.

V

A

If Scanlon's right about permissibility—if an act's being permissible consists in the fact that, according to the single-act principles that you should use in deciding whether to do it, it's false that you shouldn't do it—then we should take (28) and not (30) as owning permissibility.

What kind of moral fact, then, is the fact you should not (switch and fail to believe with justification that switching is harmless). Scanlon identifies three moral dimensions: permissibility, blame, and meaning (the significance for your relations with others of the way in which you decide how to treat them). (28) is not about permissibility, since it's not the principle you're to consider in deciding whether or not to do the act. It's not about blame since it makes a normative claim that you shouldn't be in a particular state. And I don't think it's only about meaning. A person who flicks the switch despite her belief that that's harmful does not merely reason incorrectly about whether to flip the switch; she also mistreats her neighbor by failing to "take due care."

I think we need another dimension, then. I think that it's probably *wrongness*. That's going to be a hard case to make. Here I'll just point out that Scanlon's own contractualist theory of wrongness might be recruited to make it.

B

For Scanlon an act's being wrong is its violating a principle that no one can reasonably reject given that she shares with others the aim of finding principles that no one can reasonably reject given that shares with others ... An act can violate a principle only if the principle takes a simple act as its object. Principles like that are, as I've allowed, the ones that establish permissibility. So it's credible, on this picture, that an act is wrong if and only if it's impermissible.

But are acts the only things that can be wrong? It's not clear that Scanlon is in a position to insist on that.

C

Scanlon offers his theory as explaining the reason-giving force of the consideration that an act is wrong. In some places he explains that by appealing to a basic reason that each person has to avoid acting in ways to which other people can object with respect to reasons based in their own interests and claims.

In other places he offers what I take to be a better explanation. He argues that there's positive value in the fact of several persons' doing things to and with one another in ways that are governed by each person's recognition of the reasons that each is given by the others' interests and claims. The fact, as Scanlon puts it, of mutual recognition. You should avoid doing a wrong act, on this view, exactly because doing it is incompatible with mutual recognition.

D

But here's something else that's incompatible with that: interacting with people in a way that's not appropriately governed by consideration of the reasons arising from their lives. In particular, performing a permissible act as the product of reasoning that takes bad account of those reasons. (Notice that the reasoning itself, or my holding false beliefs about reasons, is not incompatible with that. They don't by themselves make it the case that I'm not interacting with others in ways governed by the right sort of reasons. For I might yet reason my way out of them and act for the right reasons. Once I've acted on the wrong reasons, on the other hand, it's too late for mutual recognition.)

E

Scanlon's very wise strategy for finding out what wrongness is, set out early in *What we owe to each other*, is to identify wrongness with whatever property turns out to best explain our having decisive reason not to perform the acts that we judge to be wrong. Let's run with that. On the version of his explanation that I've just sketched, that property is incompatibility-with-mutual-recognition. And I've claimed that acts are not the only objects that can have it. Act-and-belief complexes in virtue of which I'm treating you in a way that's not governed by the right sort of consideration of your interests and claims are another bearer of it.

If that's right, they're wrong. Welcome to the fourth dimension.

Appendix: joint requirements

It seems to me that I sometimes think about what we should do. A naïve view takes that appearance at face value. It claims that this thinking of mine addresses the question of which several acts should be performed by several of us persons.

Say that the practical requirement

(31) We should do (I do x, you do y)

is a *joint requirement* if there are no individual requirements of the form

(32) I should do x in circumstance C

(33) You should do y in circumstance D

such that (31) consists in or holds in virtue of (32) and (33). The naïve view claims that there are some joint requirements. And it claims that these form the content of at least some of the thoughts that I express by saying that we should do something.

I

Here's some evidence for the existence of joint requirements.

You have an oar. I have an oar. If we both row, we'll cross the river to where the party is. If neither rows, we're stuck here on the boring side. If only one rows, we spin out toward the hazardous falls. Each has the same interest—survival—in crossing the river or staying on this side rather than going over the falls and the same interest—party—in crossing rather than staying on this side or going over the falls.

Consider

(34) A person should row if and only if she should believe that the other will row.

You and I each satisfy (34) if each decides not to row while giving the other reason to doubt that she intends to row. But I think that, when we satisfy (4) in this way, someone has done other than what should be done. To explain this judgment, I suppose that, beyond (34), we also face in virtue of the reasons of your and my survival and partying

(35) You and I should do (you row, I row)

When each fails to row because she expects the other to fail, each has done as she should but *we* have done other than as we should.

This evidence is not overwhelming. Does the individual rower have some way of communicating with the other by means of which she might focus both rowers' expectations and intentions on the "good equilibrium" in which each intends and believes that the other intends to row? If she does, then the judgment of practical failure—of someone's doing other than what should be done—might well condemn only our individual failures to realize that coordination by sending the relevant signals. And if not—if you're clear that there was nothing the individual rower could say, nothing else that she could do, to bring the good equilibrium about—then you might find it comfortable after all to deny that anyone has done other than what should be done. Maybe our dismay about a failure to row in that case amounts only to our judgment that being stuck on the wrong bank is sad or

bad—a bad, sad outcome. Maybe it's not the judgment that someone's made a mistake in action.

I don't insist that the latter judgment is obviously correct. The example is a foot in the door.

II

A

I'm alone in the boat now. To get it across the river takes first a stroke with the left oar and then a stroke with the right. I expect I'll be too lazy to make the right stroke when its time comes. So I don't make the left. My action satisfies

(36) I should row left if and only if I will later row right.

I should row right if and only if I've already rowed left.

But it seems not to be OK, and this appearance is explained by its violating an extended requirement

(37) I should (row left and then row right).

B

What explains the fact that an extended requirement has force here? What distinguishes my rowing predicament from, say, an afternoon of occasional chores on which I should return some DVDs to the library, and on which I should pick up some ice at the liquor store, but on which I'm not extendedly required to (drop off the DVDs and pick up the ice)?

I imagine that, for ordinary versions of these errands, the total normative case for a library run—the obtaining of reasons for and against it, and their possession of the weights they have—is independent of any facts about whether I've gotten or will get to the liquor store for ice. The total normative case for an ice stop seems to me in the same sense independent of facts about my making it to the library. Consider

(38) The acts *x* and *y* are subject to a compound requirement of the performance of (*x* and *y*) only if, for at least one of these acts, the existence or weight of at least one reason for or against that act is not independent of facts as to whether the other act will be done.

That rowing left will help me across the river counts in favor of rowing left—it counts as fact—only if I will row right. So (38) signs off on (37)'s being a genuine compound requirement. The case for rowing does not decompose.

C

The principle (38) should be checked against other examples. Beyond examples it would help to have some general explanation of how it is that the deliberative indecomposability of several objects helps to make it the case that they're subject to a compound requirement. I don't have an explanation like that; I'm not sure that one exists. But we can find some systematic support for (38) by considering the deliberative difference that compound requirements make to some of the indecomposable practical problems which (38) singles out for government by them.

Suppose that (37) is not in force, so that (36) stands as the whole practical truth about rowing. Then the to-be-doneness of rowing left depends by (36) on the will-

be-doneness of rowing right. I should, in thinking about whether to row left, take as determined either my having or my lacking an intention to row right or some other attitude thanks to which I am or am not disposed to do it. When my having that attitude gives me reason to believe I won't row right, (36) prompts me to conclude, on grounds of my having it, that I shouldn't row left.

I think that with this recommendation (36) recommends a mistake. Whether I should have my anti-right-rowing attitude depends on whether getting across the river counts in favor of acts that help to accomplish this. That issue is also a determinant of whether I should row left. I shouldn't think both that left-rowing's contribution to river crossing counts in favor of rowing left and that my attitude against rowing right after rowing left is OK. I should regard the appropriateness of that attitude as an open matter—to be settled in my further deliberation about rowing. In the first appendix I argue for this principle:

(39) If the appropriateness of holding an attitude A at t depends on matters that you should be attempting to settle in your thinking, at t and beyond, about whether to x at t + 1, then the fact that you have A at t can't help to make it the case at t that you should x at t + 1.

If it's right, the conditional single-shot principles in (36) can't be the principles that require me to row. Yet I think that I am, somehow, required to row. Until some other maxim to that effect appears, I should accept the intertemporal (37).

D

A parallel argument for the two of us. I assume that, if (34) is true of any situation, it's true of a day on which I share the boat with a rower who makes no practical or epistemic mistakes. According to (34) I should row only if you will row. Since you make no mistakes in the application of (34) to your own situation, you are not now disposed to row unless you should believe that I will row. And you should believe that I will row only if and because the current state of my beliefs and intentions is such that, in the course of any further deliberation starting from that state, I will come or continue to intend to row. It thus follows from (34) that my having attitudes in favor of or against rowing makes it respectively the case or not the case that I should row. But by (39) we should reject this consequence of (34), and so we should reject (34) itself. With (34) disqualified, (35) stands as the sole candidate for a principle instructing non-mistake-making people to row.

E

You might think that, because my rowing counts as fulfilling (35) only if you row, (35) tells me to row just in case you are going to row. If my coming to row from my acceptance of (35) passes through (34) in this way, I must not appeal to (34)'s alleged unacceptability as a reason for accepting (35).

But I doubt that this is how (35) gets us across the river. How it else can it happen? You and I might form Gilbert-and-Velleman intentions:¹

(40)(me) I intend [to row if you have intention (40)(you) and if you know that I have 40(me).]

(you) You intend [to row if I have intention (40)(me) and if I know that you have 40(you).]

When each person has her part of this syndrome, and when each knows that the other has it, we each intend to row.

I think that (35) unconditionally requires of each of us that she intend to do, in circumstance C, any act of rowing that will count, in C, as part of the set of our rowing acts. In the circumstance that consists in your having (40)(you), in my having come to row by intending to row if you have (40)(you), and in your and my knowing that the other has these attitudes, my act of rowing counts as part of your and my rowing. So by (35) I am unconditionally required to form (40)(me). So if each of us knows that the other makes no mistakes, these intentions suffice for rowing. Away we row.

Notice that no such help is available to (34). Even when you've formed (40)(you), I can satisfy (34) either by rowing on the basis of (40)(me) or by not rowing. So (34) does not require that I intend [to row if you have (40)(you).]

Coordination problem solved? Not exactly. If we can't make these intentions known—the falls are too loud, there are more rowers than fit in my earshot—or if we don't know that the others make no mistakes, the (40) intentions don't suffice for rowing, and we're still stuck.

F

Suppose that each of us accepts (35) and is rational in the sense that she tends to form the attitudes that would be required if her normative beliefs were correct. So each forms her member of (40). We can't talk or pass notes, however, and neither of us knows that the other is rational. Then we're again stranded in breach of the joint requirement. And this is embarrassed by a principle like

(41) If R is a genuine requirement, and if the actors to which it applies are rational and accept R, then they can come to satisfy R by reasoning one of whose steps includes the thought that R.

I accept (41). I don't accept that these rowers stuck on the wrong bank are rational. If my and your belief in (35) is true, then we should eventually form intentions to row. In lacking the capacity to assure the others that we've formed our parts of (40), we lack the capacity to form these unconditional intentions. And in lacking that capacity, we are not rational though each of us is rational.

The last claim presupposes that not only persons but populations of persons can be rational or irrational, that a population of rational persons can be irrational, and that population-level rationality consists not only in the members' capacities for recognizing and reasoning from normative facts on their own but also in their capacities for communication. But these are the natural assumptions to make once you've granted that we face joint requirements and that rationality is the capacity to form the attitudes we judge to be required.

III

A

You and I are walking together up the street. A stranger is moving alongside us, at the same pace and in the same direction, taking care not to bump into us. Margaret Gilbert's question is What makes it the case that you and I walk together and that you and the stranger do not?

This question is sometimes answered with the proposal that you and I walk together in virtue of a *shared intention* to walk. So what's a shared intention? How do we count as holding one?

B

Michael Bratman thinks that you and I share an intention to walk in virtue (among other facts) of each person's intending that we walk. To this proposal David Velleman once objected that I can intend that p only if I take myself to be settling the matter of whether that p by intending that p and that it's not clear that more than one of us can each settle the matter of whether we'll walk.²

As Bratman points out in reply, the settling that's required for intention can take the form of settling on the assumption of a causal background that includes other persons' intentions or dispositions. Where I know that my moving the pump handle will pump water only thanks to your disposition to open a valve, I can by intending to move the handle intend to pump water into the house.

Imagine that I intend that we walk, that you intend that we walk, that at every moment the continuation of each intention is conditional on the current existence of the other, and that all of this is common knowledge. Then, says Bratman, my continuing to intend that we walk, given my belief that and the fact that you will continue to intend that we walk conditionally on my having that intention, settles the matter of whether we'll walk. And likewise for your intention.

C

But here's an important difference between the two settlings. When by intending to pump water I settle the matter of whether water will be pumped, I have an intention that not only causes the motion of the handle and of the water but that represents itself as causing those against the background of your dispositions. As Velleman would say, the intention settles the pumping matter not only in fact but notionally.

While it might be that my intention that we walk can cause our walking, it's far from obvious that it can represent itself as causing that. I can have an intention part of whose content is that it's going to cause me to walk. But for the intention to represent itself as causing our walking, its own content must include that it causes you to walk by causing you to maintain yours.

To meet this condition on the content of the intention, it's not enough that I intend to walk and that I believe that my having this intention causes you to walk. Otherwise a person who intends to drink a foul brew and who believes that her having this intention will lead a millionaire to give her some money has an intention that she and the millionaire do the set of acts (she drinks, the millionaire pays).

Perhaps my intention that we walk includes not a belief about my causal powers but my intending [to bring it about, by having and maintaining the intention that includes this very content, that you walk.] But I doubt that I can intend [to bring things about by means of this very intention].

D

Velleman offers to step around these problems by discarding Bratman's intention-that-we-walk. Velleman thinks that a shared intention can consist in a pair of mutually referential utterances in the approximate image of (8), understood as intentions [to perform individual acts if others have the corresponding intentions.] When I say "I'll walk if you will" and you say "I'll walk if you will", and when we interpret these so that each satisfies the antecedent of the intention expressed by the other, each person's saying her thing settles the matter of her walking and causes the other person to walk. Velleman claims that the collection of these utterances is a representation of our walking, shared out between us in our component representations, that causes us to walk and that represents itself as causing that.

But all I can find here is a collection of representations of individual acts that represent themselves as causing those acts conditionally on other representations' being in place. I don't see that any of these intentions represents itself as causing anything other than the individual act. Nor do I see how the collection of those individual representations counts as a single representation of our walking, let alone one that represents itself as causing us to walk. (When you and I film each other filming each other, we make, not a composite representation of two people filming, but two representations, each of one person filming. Instead of *Two Men and a Movie Camera*, it's *Man with a Movie Camera* and *Man with a Movie Camera II*.)

So I have trouble believing that anything in this neighborhood amounts to a shared intention. A couple of examples will next suggest that, whatever you want to say about shared intention, a system of mutually conditioned intentions is neither necessary nor sufficient for acting together.

E

In *The Third Policeman* by Flann O'Brien two laborers kill an old man with a bicycle pump and hide his money by the side of the road. Each fears that the other will go back for the money or confess the crime to the police. Each decides never to let the other out of his sight. They work side by side, eat from the same dish, sleep in the same bed.

We're out in the road. Each of us wants to go home. Each also wants not to lose sight of the other. We set off. Each intends to walk down the road so long as the other intends to walk down the road. Each knows this last fact. Each walks partly so as to keep the other walking by assuring the other that he himself intends to walk.

We don't share an intention. We don't walk together. We are just two people walking, each of whom intends to walk on the condition that the other intends to walk and partly so as to sustain the other person's intending to walk. Even after we've told one another "I'll walk if you will" we still walk one by one.

F

The show was good, and I start clapping. I do this even as I'm unsure that the others agree or will clap. The crowd is enormous, and my own clap can't promote any person's clapping later on as it can't discernibly raise the overall volume of

clapping noise. (What is the sound of two hands clapping?) But I think we should clap; we should let the performer know that we liked her. I start clapping just on the chance that the others will do it too, so that I may participate in sending this signal. If the others bring their hands together for the same reason, I think it clear that, while no one intention to clap depends on any other such intention and while no one has told anyone else that she'll clap if he will, we clap together.

G

Here's a difference between our walking together up the street and O'Brien's murderers' dragging each other down the road. When I walk with you, I think that
We should walk up the street.

I think this because I think that we should eat in the restaurant at the end of the street or that we should have a talk. This judgment governs my own intention by leading me to intend to keep pace with you and to walk in the direction you're taking so that we both end up at the restaurant or so that we manage to exchange thoughts. The murderers have no corresponding judgments about what they should do. Each thinks only that

I should get home to boil some potatoes

I should stay in visual range of the other murderer.

The same structure helps make sense of clapping together. When I clap, it's not immediately because I intend to communicate my own judgment to the performers. It's instead because I think that we should clap so as to deliver a judgment shared by large numbers of us. I can respond to this requirement by clapping even if I don't expect that my own intention to clap will sustain any other person's clapping.

H

How do our responses to a joint requirement qualify us as acting together? I'm still trying to figure this out. This is what I've got so far.

Each person has an attitude of the same content: a belief that we should walk up the street. Each token of that attitude is causally indispensable for the production of the set of acts whose requiredness forms its content. So we precisely *share* something—the belief—such that each person's having her token of that thing is causally indispensable for the event that thing's about. Shared belief plays in this telling the unifying role that Bratman and Velleman assign to networks of intentions. It's an attitude in favor of the entire event of our walking such that by sharing that attitude we make the thing happen.

IV

The thesis of joint requirements raises a “missing agent” flag for some people. They can't detect an actor capable of meeting these requirements by accepting them and acting on their basis. One reply insists that several persons can compose a single actor after all. Persons make an actor when their acts, beliefs, and intentions stand in some special relation. Perhaps when the members of a club are committed to abide by a majority vote, this qualifies them as a group actor so that it's jointly required of them that they not raid other clubs or leave their campsite a mess.

On this picture joint requirements have a fairly specialized deliberative role. They tell independently constituted group actors what to do where those group actors happen to exist. As my discussion in I and II made clear, I think that's too special. I think that joint requirements have force even over populations whose members don't satisfy any special conditions of group actorhood.

Section III suggests that this joint-requirements-for-independently-constituted-group actors view has things backward in a sense. Not only do unincorporated persons fall under joint requirements. But they can succeed in acting together—they can make an actor, if you want—if each of them acts from these requirements that apply to them whether or not they come together to meet them.

V

A

I think that in my action and deliberation I should treat the similar interests of different persons similarly. Call this my belief in moral equality. I don't understand it very well. To understand it better I'll need a more definite conception of the similarity of interests and a fuller-fledged conception of interests' being treated similarly within a person's deliberation.

B

I think that on some of the occasions when I can help, in the same kind of way, either myself or another person, I should help myself and not the other. Apparently I believe that it's okay for my action and deliberation to treat my interest and that of the other person differently.

This seem inconsistent with my first thought, though the open-endedness of moral equality makes it hard to tell. It might turn out that on the correct final conception of treating similar interests similarly—on that conception on which the claim that I ought to treat similar interests similarly comes out true—I do after all treat similar interests similarly when I help only myself.

C

Suppose that each of us, in thinking what to do, tends to put extra weight on reasons rooted in her own life. There are nonetheless dimensions on which we treat the similar interests of persons similarly at the level of our entire population. For example each person's interest enters the sum of our n deliberations via the same pattern of attention to it: there's one actor putting extra weight on a reason based in it, and $n - 1$ actors who hang smaller weights on it.

To reconcile partiality and equality you might think that we deliberate and act in a pattern that qualifies us as treating similar interests similarly in this or some other respect. You might think that this fact about *us* qualifies *me* as acting consistently with moral equality notwithstanding my practice of assigning special force to my own concern.

D

Left there the suggestion is no good. From the claim that we treat like alike in some way it fails to follow that I treat like alike in any way. I am not us. I am me. For all that I've said I am only a person who puts extra weight on his interests, surrounded by a bunch of other people who put extra weight on theirs. For all that I've said the pattern in which we act is not a fact about my action that matters to its moral character.

But suppose that everyone does her part because she thinks that we should do the set. In this way we act together. Then you can after all describe me as treating like interests alike in virtue of my participation in our treating like interests alike.

E

Imagine that, for every act that I might perform as the outcome of deliberation in which I put extra weight on the reasons attached to my own interest, that act forms part of a set called for by a joint requirement enjoying strictly symmetrical support from the reasons attached to the interests of different persons. Imagine that I do this thing, not only because it's favored by the balance of partially weighed reasons for and against my individual act, but also because it's part of what we should do in virtue of reasons given by all persons' lives weighed on a par. Then the entire course of my action and of my thinking about what to do can count at once as partial and as morally egalitarian.

This sets a problem for moral construction. I'm to revise my initial judgments about what I should do, and my initial judgments about what we should do, so that individual and compound principles come out recommending the same acts and attitudes to me. The warrant for this revision is that, without it, my judgments about the individual requirements are inconsistent. One of them, the requirement that I treat like alike, condemns the others that for favoring me.

I won't try to say now how that revision goes. Its result would be a kind of totalitarianism under which everything you do is part of what we should do. But it would be a liberal totalitarianism since one of the things we should do is, as I'll now explain, to leave each one of us alone.

F

For any situation *S* in which you and I are each deciding what to do, the reconstructed principles will recommend to me and to you acts and aims whose conjunctions are also jointly required. I don't suppose that, to be jointly required under those principles, aims must count as robustly cooperative. It might be for example that in *S* we should have (I aim to win the pie-eating contest, you aim to win the pie-eating contest). But at least one form of interpersonal inconsistency seems untenable. It's presumably false that in *S* we should do (I *x*, you try by *y*'ing to get me to do something other than *x*). Then in *S* I should not try to get you to do things unless in *S* you should do them. Since *S* is characterized prior to my *y*'ing, I should not try to get you to do things unless you should do them in virtue of reasons that hold independently of my *y*'ing.

G

This is at least halfway to the principle (1) that I've defended as a major moral constraint on getting people to do things. That principle speaks, not only against

your getting people to do things unsupported by good independent reason, but also against your getting them to do things along paths that bypass their recognition of independent reasons.

By doing my part of what we should do I can treat like alike. But this works only if we act together. If my action is not part of something that we do together, then it doesn't participate in a population-level equal treatment of all persons. In III I've argued that we act together if each [does her act because she thinks that we should do the set that it includes.] Directing your conduct in a way that runs around your recognition of the independent reasons for it puts that possibility of living with others as equals by sharing collective action with them out of reach. That is another argument for (1).

Notes

1. By putting several acts or attitudes in parenthesis I emphasize that they form the compound object of the *should not* which they follow. By putting an entire verb phrase in square brackets I put the focus on the full course of action and deliberation that it describes.

2. I've helped myself here to a boring conception of freedom on which a person counts as being unfree not to x in virtue of her lacking an alternative to x'ing and as being less free not to x in virtue of her having fewer or worse alternatives to x'ing. It's not a very good conception. It stands little chance of explaining many of the central judgments that we express by talking about freedom. But no matter. The view I'm criticizing holds precisely that coercion is wrong because it makes us in this boring way less free.

3. An argument for wrongness will be assembled from what's now section V of the first appendix and what's now section V of the second.

4. Keeping my promise from III.A. When you make a target of your future conduct, locking it in the sights of your current planning, you don't act routinely or idly or otherwise acceptably unthinkingly. So the subrational episodes that threatened complication for (6) don't come up. That's why I decided to ignore them.

5. T. M. Scanlon, *Moral dimensions*, 2008. Like Scanlon I'll suppose that a central way in which some aspect C of an act x can count as intentional is for the agent to believe that C provides a reason for x'ing and for that belief to help cause her x'ing in a normal sort of way. On that assumption (23) asserts the independence of a fact about whether to do something that causes a death from the intentionalness of killing. If you disagree, then forget about intentionalness and focus instead on this question: Can an actor's believing something to be a reason for her act help to make it the case that you shouldn't do the act?

6. You might object to (24) on the independent ground that giving the drug is wrong only if the reasons belief about skiing is *decisive* in the doctor's decision to give it. If that bothers you, feel free to weaken the principle accordingly. For example:

(24') If it's the case that, if you were to come to intend to give the drug, your belief that a ski vacation is a reason to give the drug would contribute decisively to your coming to intend to give it, then you should not give it. Because the objections to (24) that I'll be considering also stand against (4') as well, and because my arguments that (25) avoids those objections will extend to an corresponding weakening of (25) like

(25') You should not give the drug as the outcome of deliberation in which the belief that the ski vacation is a reason for giving it is decisive.

I think it OK to focus on (24) and (25).

7. Margaret Gilbert, "Walking together"; J. David Velleman, "How to share an intention".

8. Michael Bratman, "Shared intention", "I intend that we J", "Modest sociality and the distinctiveness of intention"; Velleman, "How to share". My criticisms of Velleman and Bratman are too fast to do any real damage; my aim for them is mainly to introduce the shared intention view as a foil. Bratman's mature proposal includes further conditions for shared intention that save him from the counterexample of *E*, for example by requiring that each person be disposed to help the others in carrying out what they intend.

9. I took a lot of good direction from Japa Pallikathayil's unpublished *Your money or your life* and from Stephen White's unpublished "What's wrong with coercion?". In the second appendix the argument of section I is indebted to Derek Parfit, "What we together do" and to Christopher Woodard, "Group-based reasons"; the arguments of section II and III to the work of Michael Bacharach as developed by Natalie Gold and Robert Sugden; and the argument of section V to Thomas Nagel, *Equality and partiality*.

Thanks to Richard Arneson, Anne Barnhill, Selim Berker, David Bordeaux, Lindsey Chambers, Ben Chan, Jerry Cohen, Jakob Elster, Kevin Falvey, Ashley Feinsinger, Kristina Gehrman, David Goldman, Mark Greenberg, Matt Hanser, Liz Harman, Barbara Herman, Pamela Hieronymi, Louis-Philippe Hodgson, Adam Hosein, Waheed Hussain, Sean Ingham, Diane Kierce, Sari Kisilevsky, Steven Lukes, Adam Masters, Mike Martin, Ben McKean, Thi Nguyen, Japa Pallikathayil, Georgios Pavlakios, Philip Pettit, Thomas Pogge, Michael Rescorla, Ira Richardson, Matthias Risse, Andrea Sangiovanni, Tim Scanlon, Micah Schwartzman, Wally Siewert, Sophie Siow, Tom Smith, Lucas Stanczyk, Nicos Stavropoulos, Anna Stiliz, Tiffany Teeman, Sabine Tsuruda, Laura Valentini, Emmanuel Voyiakis, Gary Watson, Stephen "UCLA" White, and Andrew Williams.