

Modest Sociality and the Distinctiveness of Intention¹

Michael E. Bratman – Stanford University

For Princeton Seminar 11/08 -- 9/25/08 -- **DRAFT**

Cases of modest sociality are cases of small scale shared intentional agency in the absence of asymmetric authority relations. Examples include: our singing a duet together, but not the activities of an orchestra with a conductor; our going to NYC together, or (to use Margaret Gilbert's example²) our walking together, but not a school trip led by a teacher; our having a conversation together, but not an exchange in which a language teacher guides a novice. Modest sociality is a deep feature of our lives, something we frequently value both for its own sake and instrumentally. A human life that was not significantly embedded in such sociality would probably be unrecognizable, and certainly be impoverished. Reflection on the underlying structure of such modest sociality may also help us think about larger scale cases, such as law and/or democracy; but those are not matters I address here.

Lets work with Gilbert's wonderful example, but with a small addition. Suppose that you and I are walking together down Fifth Avenue, but there is also another person walking alongside us – call him the Stranger. Neither of us is walking together with the Stranger in a way that constitutes modest sociality. Nevertheless, all three of us proceed at roughly similar rates without bumping into each other. What distinguishes your and my relation to each other, in our modest sociality, from each of our relations to the Stranger?

Step back: what do we want from an answer to this question? I think we have at least three main concerns: conceptual, metaphysical, and normative. We seek a articulated conceptual framework that adequately supports our theorizing about modest sociality; we seek to understand what in the world constitutes such modest sociality; and we seek an understanding of the kinds of normativity – the kinds of “oughts” – that are central to modest sociality. And throughout we are interested in the relations – conceptual, metaphysical, normative – between individual agency and modest sociality.

Distinguish two initial approaches. The first is a commonsense version of the idea of Nash equilibrium in game theory. In modest sociality each is acting in pursuit of those things he values, in light of what the other is doing. What we each are doing and what we each value are out in the open – are common knowledge -- between us. And our intentional actions are in equilibrium in the sense that each sees himself as best pursuing what he values given what the other is doing. This is a model of modest sociality as equilibrium within common knowledge.

As I see it, this model is not well-suited to explain what is special about my relation to you in our modest sociality, in contrast with my relation to the Stranger. After all, the Stranger and I do not bump into each other. The relevant intentional activities of me and the Stranger may well be in equilibrium in a context of common knowledge. This indicates that we do not yet know what is distinctive of modest sociality: we have not yet captured the connections between participants that distinguish modest sociality from the coordinated concatenation of activities of me and the Stranger.

This may suggest that we turn to a second model, one that highlights the ways in which participants in modest sociality are entitled to hold each other accountable. The idea is that the interconnections characteristic of modest sociality involve mutual entitlements to hold the other accountable for playing his role in the shared activity, and mutual obligations, each to the other, that correspond to these entitlements. If we understand these entitlements and obligations as not essentially moral, we arrive at an idea that is central to Gilbert's understanding of shared agency.

As Gilbert sees it, the move from individual to shared agency involves a move to what she calls a "joint commitment".³ Gilbert does not try to provide an analysis of this idea of a joint commitment. She sees it as a basic, non-reducible idea, one that is an analogue of the idea of an individual commitment in individual agency – as when an individual agent arrives at her own decision to act. However, Gilbert still wants to make certain substantive claims about such joint commitments. She proposes, first, that when – and only when -- there is such a joint commitment there is a "plural subject".⁴ And she claims, second, that

“obligations with corresponding entitlements inhere in any joint commitment.”⁵ I will return to the idea of a “plural subject” at the end; now I want to focus on this appeal to obligations and entitlements as a central idea within the cited second approach to modest sociality.

Now, I agree that mutual obligations and entitlements are extremely common in cases of modest sociality, though I think that these will normally be familiar kinds of moral obligations – in particular, moral obligations associated with assurance, reliance, promises, and the like. But, first, I am not convinced that such obligations are essential to modest sociality; and, second, I think that there are significant resources – conceptual, metaphysical and normative – in the territory between these two initial models, and that we do well to see if there is a viable theory in that territory.

A first step is to note a range of examples that are plausibly interpreted as cases of modest sociality in the absence of mutual obligations of performance. Perhaps, to consider another variant of Gilbert’s example, we meet by accident while walking down the street, walk together (in a shared intentional activity) for a very brief time, and then go our separate ways. Perhaps we more or less spontaneously applaud in response to an obviously outstanding performance but without any prior assurances or other signals between us. Perhaps two fans of Ayn Rand sing together even though each has made it clear to the other that neither is obligated to the other to continue: “no obligations,” they each say. (Daniel Markovitz has observed that such cases are analogues of the doctrine of “employment at will” in American law.) Perhaps one of the participants coerces the other into participating in a shared activity. Perhaps we are engaged in a morally outlandish shared activity – ethnic cleansing, for example. It seems plausible to me that there are versions of each of these cases in which there is modest sociality without mutual obligation of performance.

Of course, if one of the participants in these shared activities opted out without warning, the other might well be justifiably surprised. But it is one thing to falsify a prior expectation, another to violate a prior entitlement.

Again, a participant in some such cases might try to make a claim on the other to perform. One participant in ethnic cleansing might insist that his partner not opt out without his permission; one of the accidental co-walkers might insist that the other not opt out without his permission; the master might insist that the slave, with whom he is acting, has an obligation not to opt out without his permission. But it seems to me that in some such cases such claims on the other are, to use a remark from Bernard Williams, “bluff”.⁶ The claim that such insistence is in fact grounded in an actual obligation that is owed by the other to the person making the claim is, after all, a substantive normative claim. You do not get such claims for free: they are subject to the scrutiny of normative reflection and to the need to articulate the underlying principles of obligation.⁷ And, on reflection, these substantive normative claims about obligations of each to another in some such cases seem to me strained, at best. Yet even when we reject such substantive normative claims we can see that the parties are acting together and participating in characteristic structures of modest sociality, and that this need not involve a normative illusion on the part of those parties, or of us.

That said, I do agree with Gilbert that relevant mutual obligations are common in modest sociality; but as I see it this is because assurance-based moral obligations and the like are common in such cases. And I agree that the model of equilibrium within common knowledge will not solve our problem about what is special about modest sociality. But I want to see if we can construct a model of modest sociality that is stronger than the equilibrium model but does not rest at bottom on such obligations and entitlements – though it will still need to help explain the presence and roles of such obligations and entitlements when such there be.

A first step is to say that what distinguishes you and me from you and the Stranger is that you and I share an intention to walk together -- we (you and I) intend to walk together -- but you and the Stranger do not. In modest sociality, joint activity is explained by such a shared intention; whereas no such explanation is available for the combined activity of you and the Stranger.

This does not, however, get us very far; for we do not yet know what a shared intention is, and how it connects up with joint action. Indeed, each of the two models just mooted might be converted into purported models of shared intention.

How then should we think about shared intention? Well, how should we think about the intentions of individual agents?

We make progress in understanding aspects of mind by articulating characteristic functions or roles together with associated norms. This idea lies behind what I call the planning theory of the intentions of individuals.⁸ The planning theory sees such intentions as embedded in forms of planning central to our temporally extended agency and to our associated abilities to achieve complex goals across time and inter-personally even given our cognitive limitations. The theory appeals to the guiding, coordinating, organizing roles of intentions as elements of larger – and, typically, partial, hierarchical, and future-directed – plans. Further, the theory appeals to characteristic norms of practical rationality that are associated with these planning roles, norms whose acceptance is normally operative in a planning agent's psychic economy. Primary among these norms are norms of consistency, agglomeration, means-end coherence and stability: Intentions are to be internally consistent, and consistent with one's beliefs. It should be possible to agglomerate one's various intentions into a larger intention that is consistent in these ways.⁹ Intentions engage demands to settle in a timely way on relevant means and the like – demands of means-end coherence. And intentions should tend to be stable over time in the sense that they should involve an appropriate resistance to reconsideration and to change.

These roles and norms help distinguish intentions from the ordinary desires and beliefs characteristic of simple purposive agency. There is much to say here, but let me just briefly note three important differences. First, ordinary desires are not subject to the same norms of consistency and means-end coherence. It is, after all, part of the human condition to have desires for different things that, one knows, are not co-possible. And simply desiring something does not yet put one under rational pressure to ascertain or settle on means to it. Second, ordinary

belief does not yet pose problems of means for means-end reasoning: simply believing that, given my social awkwardness, I will offend someone at the party does not pose a problem of settling on means to do that. Third, intentions play a distinctive filtering role. If I now believe p but consider an option X that would prevent p, a decision to X would not thereby be blocked – I would just need to be prepared to update my belief about p were I to intend to X. But if I now intend p and note an option Y which would prevent p, my intention would, other things equal, impose rational pressure in favor of filtering out Y from the options for my deliberation.

How exactly are these roles and norms that are distinctive of intention related to each other? Well, given these roles these norms are fitting in the sense both that general conformity to these norms is needed for a system reliably to realize these roles, and that being guided by one's acceptance of these norms is an important element in how these roles are normally realized. There is more to say, but what is important here is the general idea that these roles and norms are systematically related, and that it is by articulating them that we more deeply understand intending.

The claim is not that these roles are realized in all forms of agency. Not all agents are planning agents. But we – adult humans in a broadly modern world -- are planning agents, and this is central to the explanation of characteristic forms of cross-temporal organization in our lives. The planning theory is a theory about the nature of intentions understood as central elements in this fundamental form of temporally extended human agency. Such intentions bring with them a characteristic nexus of roles and norms. So it seems reasonable to see such intentions as distinctive elements of the psychic economy of planning agency. This is the distinctiveness of intention.

How can we extend these ideas to shared intention? Well, why do we bother with shared intentions? what fundamental roles do they play in our lives? Here there are clear analogues, in the shared case, of the coordinating, structuring, and guiding roles of intention in the individual case. In particular, it seems natural to suppose that the characteristic roles of a shared intention to J will

include the inter-personal coordination of action and planning in the pursuit of J, and the structuring of related bargaining and shared deliberation concerning how to J.

Granted, human shared agency many times engages practices of holding accountable. But I think it is natural to see this not as a defining role of shared intention – as what shared intention is for -- but as a frequent, supporting concomitant of human shared agency. The basic answer to why we bother with shared intentions is not “so as to hold each other accountable,” but rather “so as to achieve relevant forms of social coordination and organization in our relevant thought and action”. And in the pursuit of such organization, practices of accountability will frequently come to the fore.

What norms are associated with these social roles of shared intentions? Well, it seems that there will be associated norms of social agglomeration and consistency, social coherence and social stability. Roughly, it should be possible to agglomerate relevant intentions into a larger social plan that is consistent, that in a timely way adequately specifies relevant means and preliminary steps, and that is associated with appropriately stable social psychological structures. Failure to satisfy these norms will normally undermine the distinctive coordinating, guiding, and structuring roles of shared intention.

So we have structures of roles and norms both at the level of individual intention and at the level of shared intention. How are these structures related? Here I seek a construction of intentions and related attitudes of the individuals in appropriate contexts that would, when functioning in the norm-guided ways highlighted by the planning theory of intention, play the roles characteristic of shared intention and be subject to and tend to be in conformity with central social norms characteristic of shared intention. If we had such a construction we would have reason to say that this construction is shared intention, or at least one important kind of shared intention.¹⁰

Call the idea that shared intention consists in a structure of suitably inter-related attitudes of the participants in a suitable context constructivism about shared intention. We begin with the idea that shared intentions interpersonally

structure and coordinate thought and action, and that these structuring and coordinating roles involve associated norms. We then ask: will these norm-assessable social roles emerge from the norm-assessable and norm-guided functioning of appropriate attitudes of the individual participants – attitudes with appropriate contents, in appropriate contexts, and appropriately interrelated? We seek to answer this question by constructing a structure of interrelated intentions of the individuals, and norms that apply to and guide those intentions, that would induce the social-norm-assessable and social-norm-conforming social roles characteristic of shared intention. We want to show that intentions of individuals in these contexts and with these special and distinctive contents and interrelations would, insofar as they function properly and in a way that is guided by the norms of individual planning agency, play the roles of shared intention in part by issuing in thought and action that is assessable by and conforms to central norms of shared intention. And we want to show that violations of these social norms will normally be constituted by violations by one or more of the participants of associated norms of individual planning agency.

Such a constructivism builds on the planning theory of individual planning agency, a theory that sees the intentions of individuals as distinctive. This constructivism seeks to understand the step from individual planning agency to modest sociality by appeal to certain distinctive contexts, contents, and interrelations. And it aims to characterize those contexts, contents, and interrelations in terms that are, to the extent that is possible, available within the planning theory of individual agency. In this way, while highlighting the great importance of such sociality to our lives, it seeks also to articulate a deep continuity – conceptual, metaphysical, and normative -- between individual planning agency and modest sociality.

Given these theoretical ambitions – in particular, the search for such conceptual, metaphysical and normative continuities -- what is important is whether and how constructivism can provide sufficient conditions for shared intention. We can allow for the possibility of multiple constructions, each of which provides a basis for the social roles and norms characteristic of shared intention.

In the face of purported, alternative constructions, the issue will be which makes better sense of the complexities of these forms of sociality. And we can leave open the possibility that the best thing to say may be that shared intention is multiply realizable.

How should we proceed with such a construction? Let me note, in broad strokes, seven important ideas.

(i) intentions on the part of each in favor of our joint activity

A basic element in the construction of our intention to walk together will be my intention that we walk and your intention that we walk. Appeal to these intentions ensures that an intention-like commitment to our activity is at work in the practical thinking of each. Once our activity is an element in my plans, I will both face characteristic problems of means with respect to our activity, given a need for means-end coherence of my plans, and be constrained by characteristic requirements of plan consistency with respect to our activity. This explains something we need to explain, namely: the responsiveness of each to the end of the shared activity, responsiveness that is an element in the characteristic functioning of shared intention.

Granted, there are problems here. One concerns the concept of joint activity that figures in the contents of these intentions of each. If this is the concept of shared intentional activity then the appeal to these intentions of each to understand shared intention and modest sociality would seem to move in a conceptual circle. Some – including John Searle -- might see this as showing that we will need to recognize a primitive, non-reducible concept of shared intentionality or the like.¹¹ My response is different. In many cases we have available a concept of our activity that, while it does draw on ideas of individual intentional action, is neutral with respect to shared intentionality. We have, for example, a concept of walking together in a sense in which both you and the Stranger, and you and I, are walking together. We then use such relevantly neutral concepts in the contents of the intentions involved in our construction of initial cases of shared intention. We depend on the other elements of the construction to

ensure that when these intentions connect up in the right way to the group behavior there is shared intentional activity, and so modest sociality. As Christopher Kutz notes,¹² we can then try to use these initial cases at the bottom of a kind of conceptual ratcheting in which agents draw on this basic idea of shared intentionality in the content of more complex intentions that involve concepts of shared activity that are not neutral in this way. (For example: having a conversation.)

A related worry about circularity concerns the use of “we” (or “our”) within the linguistic expressions of the contents of the cited intentions of the individuals. Does this introduce a circularity? Well, in the initial cases of our construction this use of “we” will frequently be what Christian List and Philip Pettit call “the distributed ‘we’,” one that refers to each of the participants as the separate participants in the envisaged activity.¹³ But we can also avail ourselves of a concept of a group, more or less precisely specified. We can do this so long as that concept does not bring with it the very idea of shared intentionality. I might intend, say, that those of us in this part of the park – or, perhaps, a sufficient number of those of us in this part of the park -- run toward the plane that has crashed. So long as the use of “we” (or “our”) does not itself bring with it the very idea of shared intentionality there need be no vicious circle.

A third worry about (i) concerns not circularity, but rather the very idea of my intending our action. This violates the own-action condition on the content of intention according to which the subject of an intending is always the intended agent of the intended activity. And this does seem an initially plausible constraint on ordinary intending.

One reaction to this is in the spirit of work of John Searle. Searle focuses on what he calls “we-intention”.¹⁴ What he means by this is not what I mean in talking about our intention. Searle’s we-intentions are attitudes in the head of an individual, though attitudes that concern the activity of a “we”. And these Searlean we-intentions violate the own-action condition. This may be why he claims that we-intentions are not just ordinary intentions with a special content, a content that involves the activity of a “we”. We-intentions are, rather, a special intending

attitude, to be distinguished from the ordinary attitude of intending involved in individual agency. If we suppose that the ordinary attitude of intending is subject to the own-action condition, then it will be natural to see we-intentions as distinctive attitudes rather than as ordinary intentions with a special content.

In contrast, I believe that we can make sense of ordinary intentions on the part of each of us in favor of our activity – though these will be intentions that we act, not intentions to act. We can do this by supposing that, roughly, it is common knowledge among us that we each believe

(a) that we each intend the joint activity,

(b) that for each of us the persistence of his own intention in favor of the joint activity depends causally and rationally on his continued knowledge that the other also so intends, and

(c) that if, but only if, we do both so intend then as a result we will so act.¹⁵

In believing (a)-(c) each believes that his intention appropriately leads to – and, in a sense, controls -- our joint action in part by way of its support of the persistence of the other's relevant intention. Each believes this while also believing that the corresponding intention of the other participant also appropriately leads to – and, in a sense, controls -- the joint action. In this way each can coherently intend that we act, and recognize that the other also so intends, even though this intention violates the own-action condition.

One advantage of this strategy is that it allows us to draw directly on what we know about the nature of ordinary intention, whereas this is apparently blocked by Searle's strategy, since Searle's we-intentions are not themselves ordinary intentions. In particular, the present strategy allows us to locate both intending to act and intending that we act in largely similar ways within the nexus of roles and norms highlighted by the planning theory of intention.

I conclude that condition **(i)** is coherent, need not involve circularity, and can play a basic role in our construction of shared intention. So let me turn to other elements of this construction.

(ii) interlocking intentions

The idea is that the intentions of each are also to interlock in the sense that each intends that the joint activity go in part by way of the relevant intentions of each of the other participants. I intend that we walk in part by way of your intention that we walk; and vice versa. There is a semantic interconnection between our intentions in favor of the joint action.

This condition should be distinguished from one that says that each believes that the other's intention will be effective. It is the content of, in particular, my intention that includes a reference to the role of your intention, and vice versa. And intentions are distinctive.

This interlocking of our intentions goes beyond the idea in (i) that we each intend the joint activity. To see this, suppose that you and I each intend that we go for a walk by twisting the other's arm and forcing him to walk in a way that does not involve any intention on his part that we walk together. Each might assert in, as it were, the mafia sense, that he intends that we go for a walk; and each might say to the other, somewhat ominously, "we are going for a walk". In such a mafia case we have (i). But we do not have (ii). And that is one reason why there is no shared intention to walk together.

(iii) intentions in favor of meshing sub-plans

There can be cases in which each of us intends the joint activity in part by way of the corresponding intention of the other, and yet one or both of us intends to side-step or override, perhaps using deception, the sub-plans of the other. To change examples: perhaps I intend that we go to NYC in part by way of your intention that we go, but I intend to trick you into taking the train despite your intention to drive. I do not intend that we go to NYC by way of sub-plans that are jointly compatible. But it seems that in shared intention there will be, in contrast, a commitment to mutual compatibility of the relevant sub-plans of each. (Here I bracket complications raised by cases of competitive shared activities.)

This commitment to mutual compatibility of relevant sub-plans helps explain the coordinating role of shared intention. If you intend that we go to NYC by driving, and I intend that we go by train, we have a problem. If we share an

intention to go to NYC, then we will normally try to resolve that problem by making adjustments in one or both of these sub-plans, perhaps by way of bargaining, in the direction of co-possibility. We want our construction to account for this standard norm-conforming functioning of the shared intention. And a natural way to do that is to use in the construction the idea that each not only intends the joint activity, but also intends that this joint activity proceed by way of sub-plans of the participants that mesh in the sense that they are co-realizable and so can be consistently agglomerated. If each of us intends that our sub-plans for going together to NYC mesh, then we will each be under rational pressure appropriately to solve the problem posed by our conflicting sub-plans.

(iv) disposition to help if needed

Suppose – as in (i)-(iii) – that I intend that we walk together by way of your corresponding intention and meshing sub-plans. My intention does not see your contribution to our joint activity as merely an expected pre-condition of our going, a pre-condition to which I am, as Nicholas Bardsley puts it, simply “adding-on” and “providing the finishing touch”.¹⁶ Rather, your contribution to our walking is itself a part of what I intend: part of what I intend is that we both walk, where that involves your walking in part by way of your intention that we walk.

This means that the demands of means-end coherence and of consistency apply to my intention in favor of, inter alia, your playing your role in our joint activity: I am under rational pressure in favor of means to that, and in favor of filtering options incompatible with that. This will normally mean that, insofar as I am rational, I will be to some extent disposed to help you play your role in our walking together if my help were to be both needed and not too burdensome on me.

Granted, there can be attenuated cases in which I intend our acting together in a certain way and am fully confident that you will not need my help, but also see any form of help I might offer as too burdensome. It is possible that in such attenuated cases there are intentions of the sort cited in (i)-(iii) and yet the absence of a disposition on the part of each to help the other if needed. But I think

that these will be attenuated cases of intending the joint activity, and that in central cases of shared intention the intentions in favor of the joint activity will induce at least a minimal disposition to help if needed. So, given our concern with robust sufficient conditions for modest sociality, I will suppose that in the central cases of interest the intentions in (i)-(iii) are not attenuated in this way. These non-attenuated intentions will involve, by way of rational pressures toward means-end coherence, at least a minimal disposition to help the other if need be.¹⁷

(v) interdependence in the persistence of each person's relevant intentions.

In defending the idea of each individual participant's intending the joint action I appealed to each participant's belief in the interdependence of the persistence of each participant's relevant intentions. In a standard case of shared intention, then, such beliefs about interdependence are brought on board by way of each person's intending the joint action. Given that we seek robust sufficient conditions for modest sociality we can add that these beliefs are indeed true and the participants are not proceeding in error: there is in fact this interdependence in persistence. This does not require – what would be puzzling -- that the onset of each participant's intention depends on the onset of the other's intention. It requires only interdependence in persistence.

(vi) Joint-action-tracking mutual responsiveness in the intentions and actions of each.

This is a condition on the connection between shared intention and joint action in modest sociality. Once this condition is recognized, however, it will frequently find its way, at least implicitly, into the contents of the intentions of the participants. The idea is that the standard route from our shared intention to our joint action will go by way of mutual responsiveness of each to each in a way that tracks the joint action. There will be responsiveness of each to each both in relevant subsidiary intentions and in relevant actions. And both forms of responsiveness track the

intended joint action: this is not the mutual responsiveness of two soldiers who are fighting each other to the death.

In saying this I am presupposing that the connections between your intentions to act and your actions are appropriate for the individual intentionality of your actions; and similarly concerning the intentionality of my actions. My question is: what is normally present, in the connection between our intentions and our joint action, in a case of modest sociality, that goes beyond these connections to our individual intentional actions? And my answer is: responsiveness of each to each both in intention and in action, a responsiveness that tracks the intended joint action.

Finally,

(vii) Common knowledge among the participants of all these conditions.

In shared intention the participants will normally know of the fact of the shared intention. Such epistemic access to the shared intention will normally be involved in further thought that is characteristic of shared intention, as when we plan together how to carry out our shared intention. Since such shared planning is part of the normal functioning of the shared intention, we need an element in our construction of shared intention whose functioning involves such thoughts and knowledge of each about our shared intention.

Here something like a common knowledge condition seems apt, at least given our primary interest in sufficient conditions.¹⁸ It is common knowledge amongst us that p if we each know that p , and both p and the fact that we each know it is out in the open amongst us. So we are each at least in a position to know that the other knows, to know that the other knows that we know, and so on. How more precisely to understand these ideas is a difficult question; here I simply work with an intuitive idea of common knowledge.

There is more to say; but you get the idea. The central claim is that these interdependent and interlocking intentions of the individual participants, all in a context of common knowledge, will, in responding to the rational pressures specified by the planning theory of individual agency, function together in ways

characteristic of shared intention. This structure will, when functioning properly (in the sense of the planning theory), normally support and guide coordinated social action and planning, and frame relevant bargaining and shared deliberation, in support of the intended shared activity. And conformity to characteristic social norms of social agglomeration, consistency, coherence, and stability will emerge from the norm-guided functioning of these interrelated attitudes of the individuals.

While I cannot defend this claim fully here, let me highlight one central element in its defense. Built into (i)-(iii) is the idea that when, in a basic case, we intend to J, I intend that we J in part by way of your analogous intention and meshing sub-plans. This complex content of my intention connects it with your intentions and thereby imposes rational pressure on me, as time goes by, to fill in my sub-plans in ways that, in particular, fit with and support yours as you fill in your sub-plans. This pressure derives from the rational demand on me – a demand that comes from the planning theory of the intentions of individuals -- to make my own plans means-end coherent and consistent, given the ways in which your intentions enter into the content of my intentions. By requiring that my intention both interlock with yours, and involve a commitment to mesh with yours, the theory ensures that rational pressures on me to be responsive to and to coordinate with you – rational pressures characteristic of shared intention -- are built right into my own plans, given their special content and given demands of consistency and coherence directly on my own plans. And similarly with you. So there will normally be the kind of mutual, rational responsiveness in intention -- in the direction of social agglomeration, social consistency, and social coherence -- that is characteristic of shared agency.

This point is fundamental. The theory works, in part, by building appropriate reference to the other directly into the contents of the intentions of each. It does not merely appeal to an unarticulated “sense of the other” in something like what John Searle calls the “Background”.¹⁹ The theory seeks, rather, to generate much of the relevant normativity at the social level out of the individualistic normativity that is tied primarily to the contents of the intentions of each. The theory seeks contents of the intentions of each that ensure, given rational demands on those

intentions of each (demands rooted in the planning theory of individual agency), responsiveness to central social demands. The ideas of intending the joint activity, of interlocking, and of intended mesh are at the heart of this procedure for providing an individualistic source for assessability by and conformity to basic social norms that are central to shared intention and modest sociality.

Constructivism about shared intention begins with an underlying model of individual planning agency. This model highlights roles and norms characteristic of individual intending and planning. Constructivism then seeks a conceptual, metaphysical, and normative bridge from such individual planning agency to modest forms of sociality. This bridge draws on ideas of intention, content, interlocking, mesh, mutual responsiveness, interdependence, and common knowledge. It applies the norms of the planning theory of individuals to the distinctive intentions and plans of individuals that are constructed using these further ideas. In these ways this bridge tries to articulate individualistic infrastructures that, when functioning properly, realize significant forms of social-norm-conforming and social-norm-assessable social functioning.

As noted, Gilbert's alternative strategy is to see non-moral obligations of each to another as partly constitutive of shared agency. In my judgment, however, this too quickly passes over complex bridging structures of the sort that I have been highlighting. Though modest sociality does normally involve familiar forms of obligation, when we examine such sociality under our philosophical microscopes we can discern, at bottom, even more fundamental intentional, causal, semantic, epistemic, and rational structures.

Return to the model of equilibrium within common knowledge. You and the Stranger may be acting in ways that are in equilibrium in a context of common knowledge. Each knows what the other intends, knows, and does; each acts in part in the light of this knowledge of the other; and each knows that if both so act there will be a coordinated concatenation of their walking actions. How does this fall short of the conditions outlined in (i)-(vii)?²⁰

Well, cases of equilibrium within common knowledge involve something close to the interdependence cited in condition (v). But I do not think that such equilibrium within common knowledge ensures any of the conditions (i)-(iv) or (vi).

This is largely because of the distinctiveness of intention: to believe or know need not be to intend.²¹ In particular, though each believes that there will be the cited coordinated concatenation of walking actions it does not follow that each intends that. To intend the coordinated concatenation each would need to be disposed to take it as an end for his own means-end reasoning, to be guided in action by this end, and to filter potential options for deliberation with an eye to their compatibility with this end. But this may not be true of you and the Stranger even though your actions are in equilibrium within common knowledge. Perhaps the Stranger is looking for ways to thwart your progress down the street without physically interfering with you, even though he sees that you are indeed progressing down the street and he is doing what he thinks best given that you are. (Such is life in NYC.) This Stranger does not intend that the two of you walk together down the street, or that your sub-plans mesh. Given what he knows to be the limits on his powers, he does expect that you will in fact walk in the way you are walking. And he intends to respond to that, and so expects that there will be a coordinated concatenation of the walking actions of each. But this is not yet to intend that coordinated concatenation.

Nor need the Stranger be disposed to help you if you need it. And his responsiveness to you may not track just the coordinated concatenation but rather a more complex end that also includes thwarting you if, contrary to his expectation, it turns out that he can.

Again, perhaps the Stranger expects that your walking will be the issue of your relevant intentions and yet does not intend that. He may prefer a mechanism that would issue in your walking in a way that bypasses your intentions – perhaps, as in a mafia case, by pushing you along. Being a realist and not especially strong, however, he does not believe that this is what will happen; so he expects that you will walk by way of your relevant intention, and he does what he sees as

best given that. But he does not intend that your intention be efficacious. So his intentions do not interlock with yours.

Equilibrium within common knowledge does not, then, ensure any of conditions (i)-(iv) or (vi). That is why such equilibrium does not ensure modest sociality.

Constructivism about shared intention does, then, provide a path between the two models with which we began, a path that responds to our conceptual normative, and metaphysical concerns. While constructivism goes beyond the equilibrium within common knowledge approach, it remains conceptually conservative in the sense that the concepts it uses – with the possible exception of the concept of common knowledge – are available within the theory of individual planning agency. In this respect it stops short of Gilbert's introduction of a new, non-reducible idea of a "joint commitment". Constructivism is also conservative in its approach to the normativity of modest sociality. It sees the applicability of and conformity to characteristic social norms (e.g., norms of social consistency and coherence) as emerging from the guidance of the relevant intentions of the individual participants by the norms of individual planning agency -- though it allows both that the participants may also go on to internalize these social norms, and that familiar moral norms of mutual obligation are commonly engaged.

What about the metaphysics of modest sociality? The basic picture is a metaphysics of interconnected, individual planning agents. Given these interconnections there will normally be a systematic impact on the world of this overall structure of interconnected planning agents. So it will frequently make sense to see this overall structure as a cause of these outputs, and in that (modest) sense as a group "causal agent."²² In appealing to a group causal agent the thought is not that such causation is metaphysically distinct from the underlying causal processes involving the organized elements of the group. That the group is a causal agent consists in facts about the relevant underlying structure of individuals and casual processes involving those individuals.

Is such a group causal agent the subject of relevant shared intentions? I think that in modest sociality the answer will normally be "no". To talk of a subject

who intends is to see that subject as a center of a more or less coherent mental web. The idea of a subject who intends X but has few other intentional attitudes seems a mistake. This is a minimal lesson of Donald Davidson's work on the holism of the mental. But in cases of modest sociality the sharing will typically be partial and limited. This need not block the idea of a group causal agent that can be limited in its causal impacts and, perhaps, quite temporary. But the minimal holism of the mental distinguishes this idea of a causal agent of a shared intentional activity from the idea of a subject of shared intention. Shared intentional activity normally involves shared intention. But we should not expect that in modest sociality there will in general be a sufficiently robust, coherent web of relevant shared attitudes to give substance to the idea that the group is the subject of the shared intention. In this respect constructivism diverges from Gilbert's view, noted earlier, that in shared intentional activity there always is a "plural subject" that is literally the subject of the involved shared intentions.²³

Constructivism about shared intention builds on the planning theory of individual planning agency. Insofar as this constructivism is successful it helps support a general point about the significance of planning structures. The general point is that the introduction of these planning structures within our theory of individual agency puts us in a position to articulate theoretically rich and plausible positions in the middle of territory characterized by central debates in the philosophy of action. This is what has happened here, in our investigation of modest sociality. And elsewhere I have argued for a somewhat analogous view about individual self-governing agency.²⁴ The planning theory helps us model strong forms of agency, individual and shared, while drawing only on modest conceptual, normative, and metaphysical resources – though resources that do go beyond a simple desire-belief model of agency. And this supports the conjecture that our planning capacities are at the bottom of two further, fundamental practical capacities: our capacity for self-governance and, as I have argued here, our capacity for modest sociality.

¹ This is an overview some of the ideas in a forth-existing and as yet untitled monograph on shared agency. Some parts of this essay are drawn from my essay "Shared Agency,"

forthcoming in Chris Mantzavinos, ed., Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice (Cambridge University Press). I thank Cambridge University Press for their permission to use this material here. Earlier versions of many of these ideas come, in turn, from a quartet of essays of mine on this subject reprinted in my Faces of Intention: Selected Essays on Intention and Agency (Cambridge University Press, 1999). Earlier versions of the present essay were presented at the Pacific APA, March 2008, the Bristol Workshop on Joint Agency, July 2008, and Rutgers University, September 2008. I have benefited from the discussions on those occasions – in particular from the prepared comments of Abraham Roth at the APA session. My thinking about these issues has benefited over the years from numerous interactions with Margaret Gilbert and, more recently, from interactions with Scott Shapiro and Facundo Alonso. Work on this essay was supported in part by a Fellowship at the Stanford University Humanities Center.

² Margaret Gilbert, “Walking Together: A Paradigmatic Social Phenomenon,” Midwest Studies 15 (1990): 1-14.

³ See her “What Is It for Us to Intend?” in her Sociality and Responsibility (Lanham, MD: Rowman & Littlefield, 2000), 14-36.

⁴ “What Is It for Us to Intend?” pp. 19, 22.

⁵ “What Is It for Us to Intend?” p. 25. Abraham Sesshu Roth develops an importantly different version of this idea; but here I focus on Gilbert’s version. See Abraham Sesshu Roth “Shared Agency and Contralateral Commitments,” The Philosophical Review (2004): 359-410.

⁶ Bernard Williams, “Internal and External Reasons,” in his Moral Luck (Cambridge University Press, 1981): 101-113, at 111.

⁷ This point has also been insightfully emphasized by Facundo Alonso under the guise of what he calls “Hume’s challenge”. See Facundo Alonso, Shared Intention, Reliance, and Interpersonal Obligation (Stanford University Ph.D. Thesis, 2008).

⁸ See my Intention, Plans, and Practical Reason (Harvard University Press, 1987; reissued by CSLI Publications, 1999), and my Faces of Intention.

⁹ In formulating the agglomeration principle in this way I have benefited from Gideon Yaffe, “Trying, Intending and Attempted Crimes,” Philosophical Topics 32 (2004): 505-532, at 510-522.

¹⁰ For some complexities see my “I Intend that We J” as reprinted in my Faces of Intention, at pp. 143-4.

¹¹ See John Searle, “Collective Intentions and Actions,” in P. Cohen, J. Morgan, and M.E. Pollack, eds., Intentions in Communication (Cambridge: MIT Press, 1990): 401-415.

¹² Christopher Kutz, Complicity: Ethics and Law for a Collective Age (Cambridge University Press, 2000), at 86-88.

¹³ Christian List and Philip Pettit, Responsibility Incorporated (unpublished manuscript).

¹⁴ John Searle, “Collective Intentions and Actions”.

¹⁵ See my “I Intend that We J.” Given our interest in robust sufficient conditions, I include here a common knowledge condition. I say a bit more below about the very idea of common knowledge; but I do not try to discuss the large literature on this idea or to provide an analysis of this idea. Further, I will not try to discuss whether such common knowledge is necessary for each to intend that we so act. However, below I do point to a reason for an even broader common knowledge condition in our construction.

¹⁶ Nicholas Bardsley, “On Collective Intentions: Collective Action in Economics and Philosophy,” Synthese (2007): 141-159, at 145.

¹⁷ In this paragraph I have benefited from discussion with Facundo Alonso.

¹⁸ And recall that we have already appealed to a similar, though more limited, condition in our effort to explain the idea, in (i), of my intending that we J.

¹⁹ John Searle, "Collective Intentions and Actions," at 413-15.

²⁰ A question that is highlighted by Natalie Gold and Robert Sugden, "Collective Intentions and Team Agency," The Journal of Philosophy (2007): 109-137.

²¹ Gold and Sugden remark that "Bratman might object that it is too glib to interpret P1's intending that J come about 'because of' P2's intention as the idea that P1 believes that P2 has the corresponding intention and acts on the basis of this belief." ("Collective Intentions and Team Agency," 115) On my view, this identification of intending with belief is, quite simply, false.

²² I am agreeing here with Bjorn Peterrson, who credits Time Crane for introducing the term "causal agent". However, I do not agree with Peterrson that this poses a problem for my theory. See his "Collectivity and Circularity," The Journal of Philosophy (2007):138-156, esp. p. 148.

²³ Philip Pettit and David Schweikard arrive at a similar conclusion about Gilbert's idea of a "plural subject" in their "Joint Actions and Group Agents," Philosophy of the Social Sciences 36 (2006): 18-39, at 32.

²⁴ In my Structures of Agency: Essays (New York: Oxford University Press, 2007) I sketch an approach to self-governance that draws substantially from the planning theory, and provides another aspect of the fecundity of planning structures. And this approach to self-governance is itself a kind of intermediate position between a Humean desire-belief model and a broadly Kantian theory that appeals to forms of willing that are not locatable in the natural order.